EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L10	24	(SLA (service adj level adj agreement)) and L7	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 08:05
L9	8	(SLA (service adj level adj agreement)) and L8	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 08:05
L8	73	L6 and L2	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 08:05
L4	36	(SLA (service adj level adj agreement)) and L2	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 08:05
L3	0	(SLA (service adj level adj agreement)) same L2	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 08:05
L7	355	L5 and L2	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 08:04
L6	55695	"709"/$.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 08:02

L5	112538	"370"/$.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 08:02
S17 7	11	(SLA (service adj level adj agreement)) same (replacement near6 service)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 07:33
L2	2098	optical near5 router same switch$5	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/11 07:33
S17 6	104	(SLA (service adj level adj agreement))and (replacement near6 service)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/10 15:56
S17 5	0	(SLA (service adj level adj agreement))and (negotiate near4 replacement near6 service)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/10 15:55
S17 4	0	(SLA (service adj level adj agreement))and (negotiate near4 replacement near6 service)	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/10 15:55
S17 3	0	((SLA (service adj level adj agreement)) near5 (manag$5 monitor))and (negotiate near4 replacement near6 service)	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/10 15:55

S42	3	((SLA (service adj level adj agreement)) near5 (manag$5 monitor))and (optical near3 network) and (statistic$5 and pattern$5 and trend$5 and penalty and negotiat$5 and billing and accounting)	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/10 15:54
S17 2	0	09/930375	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/10 15:22
S17 1	2	"6366563".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2007/07/10 15:22

IN THE UNITED STATES PATENT & TRADEMARK OFFICE
PROVISIONAL APPLICATION FOR PATENT COVER SHEET

To: Assistant Commissioner for Patents
Box Patent Application
Washington, D.C. 20231, USA

Docket No.: 12630ROUS01P

Inventor(s): Waichi C. LO et al.

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR §1.53(c).
The following are enclosed:

☒ Specification — Pages: 61

☒ Drawings — Sheets: 7

☐ Assignment Papers (cover sheet(s) and document(s)). Please record and return to the undersigned.

☒ Return Receipt Postcard

☐ Other: ____

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government: ☒ No

TITLE OF THE INVENTION	
CONTROLLER FOR AN OPTICAL NETWORK	
INVENTOR(S) NAME(S)	**RESIDENCE**
Waichi C. LO	169 Hemlo Crescent, Kanata, Ontario K2T 1E5 CANADA
Michel M. DALLAIRE	1276 Perry, Aylmer, Quebec J9H 5C9 CANADA
Christopher L. HOBBS	6-310 Cathcart Street, Ottawa, Ontario K1N 5C4 CANADA

☐ Additional inventor(s) are named on a separately numbered sheet attached hereto.

FEES: Basic Fee: $150.00
Assignment(s): ____ x $40.00 =

TOTAL FEE: $150.00

The Assistant Commissioner is hereby authorized to charge the TOTAL FEE above, and to charge any additional fees which may be required or credit any overpayment, to **Deposit Account No. 14-1315.**

CORRESPONDENCE ADDRESS:

NORTEL NETWORKS LIMITED
Intellectual Property Law Group
P.O. Box 3511, Station "C"
Ottawa, Ontario, Canada K1Y 4H7
Phone: 613-768-3009
Fax: 613-768-3017
Date: May 31, 1999

Yours very truly
Waichi C. LO et al.

By Aprilia U. Diaconescu,
Patent Agent
Registration No. 37,989

CONTROLLER FOR AN OPTICAL NETWORK

The OPTera Network Controller (ONC) is a software component of an OPTera Packet Solution (OPS) network.

1.1 The OPTera Packet Solution

The OPS is a family of products that provides a carrier with a network that integrates the optical and packet layers into one infrastructure that carries Internet Protocol (IP), asynchronous transfer mode (ATM), frame relay (FR), time-division multiplexing (TDM), and SONET/SDH traffic.

The OPS offers the following features:

- intelligent dynamic network configuration that enables integration of the packet and transport layers and the ability to share network topology and state information between these two layers and optimize resources accordingly.
- packet-layer unification - the ability to consolidate IP, ATM, frame relay and TDM networks into a single core network.
- unified network management - a single interface to all OPS components.
- scalability - the ability for the packet layer of the network to scale from gigabit capacity to hundreds of terabits.
- 99.999% reliability - the above features enable system and network availability 99.999% of the time.

To be considered an OPS, a network must provide intelligent dynamic control of the network using an OPTera Network Controller (ONC), which is the software that provides this control, and one or more of the following components, whether they are the OPTera products listed below or any other non-OPTera equivalent.

- OPTera Packet Core (OPC) - the core router within an OPTera Packet Solution network, or any other core router
- OPTera IP Access Device - a router on the edge of an OPS network for allowing IP traffic access to the core, or other access device
- OPTera Multiservice Switch - a switch on the edge of an OPS network for

processing IP, ATM, and frame relay traffic.

- OPTera Transport Switch - the OPTera Connect, an optical layer switch.
- OPTera Network Management - offers a unified operations, administration, maintenance, and provisioning interface for all of the components of the OPTera Packet Solution.

1.2 The OPTera Network Controller

The purpose of the ONC is to maximize a carrier's revenue by distributing and optimizing network resources to increase the volume of traffic that can flow through a network and/or to prioritize the highest revenue-generating traffic.

The ONC:

- collects data about network traffic and topology by querying the layer 3 router (OPC), and by accessing information about layer 1 connections.
- is constrained by network policies. The policies are rules configured by the carrier's human network operator. The policies reside on a server that the ONC queries regularly.
- monitors network conditions such as:
  - failure conditions; for example, link failure, and node failure (OPC failure)
  - network congestion. Each carrier determines its own congestion thresholds. A carrier will have different threshold values for different quality of service agreements. The human network operator enters the threshold information into the network policies.
  - a request for additional bandwidth
  - link under-utilization
- Increases the use of optical resources by defragmenting the optical network
- balances network traffic in one of two ways:
  - automatically, by redefining layer 3 paths or by changing layer 1 connectivity
  - by sending its recommendations to the human network operator, who can then decide whether to implement the recommendations.

2.1 Product Requirements

The architecture specified in this document meets the following product requirements. The ONC must:

- support up to 64 nodes within a carrier's administrative zone
- base its actions in the network on policies
- perform network load balancing under the following conditions
  - — fault and recovery conditions
  - — service requests that cannot be handled by access devices closer to the edge of the network
  - — installation and decommissioning of equipment
  - — detection of congestion and high traffic in the network
- use multiprotocol label switching (MPLS), a technology for speeding up network traffic flow by setting up a specific path for a given sequence of packets.
- support Unified Label-Switched Paths—paths between OPTera packet switches. Although not a requirement, the ONC will also support standard MPLS label-switched paths (paths between non-Nortel label-switched routers)
- understand and preserve quality of service (QoS) and class of service (CoS) based paths
- support MPLS path redundancy (for fast path restoration) when performing load balancing; that is, the ONC must be aware of existing redundant paths (paths with the same source and destination), and set up a new redundant path each time it creates a new MPLS path.
- include a management interface for the human network operator to access information about fault conditions, configuration and security
- report traffic engineering information to the human network operator regarding what the ONC has discovered about network conditions, and a record of the ONC's actions
- manage transport bandwidth required by packet services; that is, map packet layer requirements to optical layer resources
- perform defragmentation of the optical layer
- be independent of an operating system platform (enough to ensure portability on platforms used by the OPTera Packet Solution components)

The architecture will be expanded in later releases of the ONC to meet the following product requirements:

- flexible service level agreements. This means the ONC can negotiate a trade-off of QoS for a lower service price to mitigate a bandwidth scarcity, or to accommodate higher priority, higher revenue traffic.
- the capability to predict network congestion, and to balance traffic in response, before congestion thresholds are reached
- the capability to permit link de-loading when a link needs to be taken out for service or repair
- an open interface to peer ONCs in the administrative zones of other carriers

The following requirements are not addressed in this document:

- the ability to plan for scheduled services, such as regularly scheduled events requiring more bandwidth
- the ability to permit reserved bandwidth time sharing among two or more carriers

The following is outside the scope of this product:

- The use of the ONC as an off-line network planning tool

2.2 Architectural Requirements

The ONC has the following architectural requirements. These requirements must be adhered to when the architecture is extended.

- The ONC is a partitioned and layered collection of software entities. This is essential to achieve the expected expansion in ONC functionality.
- The ONC software is decoupled from the actual operating system on which it is running by means of an adaptation layer. This adaptation layer is a standard operating system interface that allows the ONC to run on multiple platforms.
- The infrastructure of the ONC is independent of the actual algorithms used for the traffic engineering computations. The algorithms can be replaced without modifying the infrastructure.
- The communications protocols used within and between ONCs are standard, pre-defined interfaces offered by the packet and transport layers. The ONC avoids defining new interfaces.
- The ONCs and their components require minimum manual intervention and datafill. Even acting upon ONC recommendations manually requires

less manual intervention than managing a network without an ONC.

- ONCs make use of the existing infrastructure of the devices they are controlling; for example, at layer 3, ONCs are addressed by IP addresses because that is how routers are addressed.
- Although the ONC is not designed to be an offline network planning tool, ONC algorithms and rules can be used to create such a tool.

2.3 Performance Requirements

The ONC must meet the following performance requirements:

- *x* transactions per second or hour
- *x* messages per second
- *x* path requests per second or hour
- *x* queries per second

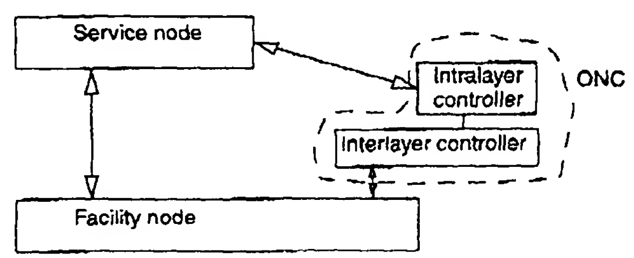
3.1 ONC Conceptual View

The OPTera Network Controller (ONC) provides a network with intelligent dynamic resource management between a service node and a facility node. A service node provides a path-oriented service, is managed by a service level agreement and engineers network resources to meet the agreement at the optimal cost. A facility node is a lower level resource required by the service node. The ONC's functions can be divided into two categories:

- Intralayer controller - at the service node layer, the ONC understands the service level agreement and manages resources at this layer.
- Interlayer controller - the ONC allocates resources from the facility node to be used by the service node.

Figure 1 illustrates the ONC conceptual view.

Figure 1: ONC Conceptual View



3.2 Implementation Example

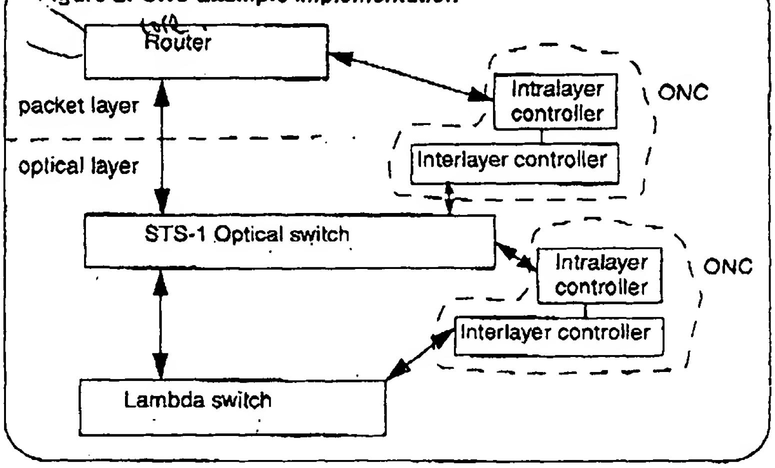
Figure 2 illustrates one example of how the ONC could be implemented. In this example, the ONC provides resource management:

- at the packet layer by balancing traffic, adjusting the bandwidth of MPLS paths, and automatically setting up end-to-end paths
- at the optical layer by defragmenting the optical network, thereby allowing the network to make use of previously stranded resources
- between the optical and packet layers, by allowing optical resources to be used directly by the packet layer to respond to congestion or increased

demand at the packet layer.

- at different layers within the optical layer; for example, between an optical switch and a lambda switch by allowing the optical switch to use the resources of the lambda switch.

Figure 2: ONC Example Implementation



3.3 Network View

The ONC is portable software designed to run on any operating system. In an OPS network, however, an ONC resides on an OPS processor shelf at each node.

When the ONC is active in a network:

- it gains a global view of the network through:
  - — automatic discovery of network topology at layer 3, and information about the location of layer 1 connections
  - — accessing data regarding network traffic levels and status of the network
- it is controlled by means of user-defined policies (rules configured by the human network operator). By adjusting the policies, the network operator can customize the ONC's actions.
- it facilitates interworking between the optical and packet layers. Specifically, the ONC maps packet layer service requirements to available

optical layer resources by:

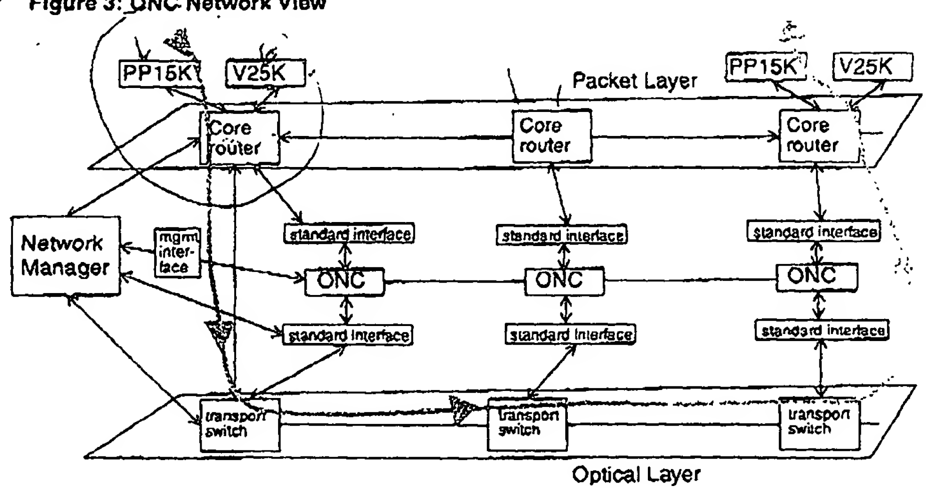
— dynamically setting up, tearing down or altering connections at the layer 1 or layer 3 level.

— performing load balancing of traffic on MPLS paths

- it records an audit trail of every action performed.
- it explains its actions in terms comprehensible to the network operator.
- it makes only revenue-positive or revenue-neutral changes to the network. This implies that, if a condition occurs that the ONC has not been programmed to recognize, any action the ONC takes will not make matters worse.

Figure 3 illustrates the network view of the ONC. The grey line represents a traffic flow and its progression through the network. In this diagram, traffic enters the network through a multiservice switch, which sends the traffic to the core router. The core router sends the traffic to a transport switch. The ONC has already determined that this transport switch is the optimum connection at the optical layer and has requested this connection through a standard interface to the optical layer. The traffic proceeds along the optical layer and is forwarded to its destination core router and out through a multiservice switch. While the traffic is flowing from its source to destination, the ONC constantly monitors the traffic and ensures that the network topology is optimized.

Figure 3: ONC Network View



3.4 ONC Efficiency

In any network there is unused bandwidth that is inaccessible because of:

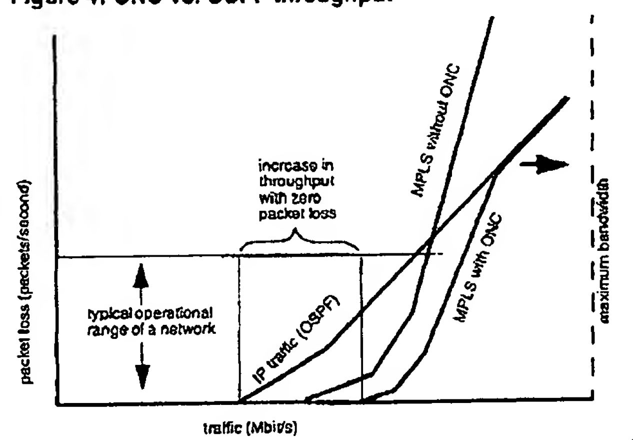
- redundant paths that are set up in case of link failure where these paths are either not being used or not used efficiently for low-priority traffic
- the nature of traffic flows; that is, a flow of traffic along a path will never use 100% of the available bandwidth on that path
- bandwidth that is committed, or reserved, to fulfill a quality of service agreement

The ONC enables the network to access some of this unused bandwidth by:

- dynamically setting up MPLS paths that:
  - — reduce the amount of bandwidth required for protection (see Figure 5)
  - — balance traffic and make optimum use of bandwidth
  - — reduce the amount of bandwidth required to be reserved for QoS
- establishing new connections at layer 1 or changing connections at layer 1 to adapt to changing traffic patterns

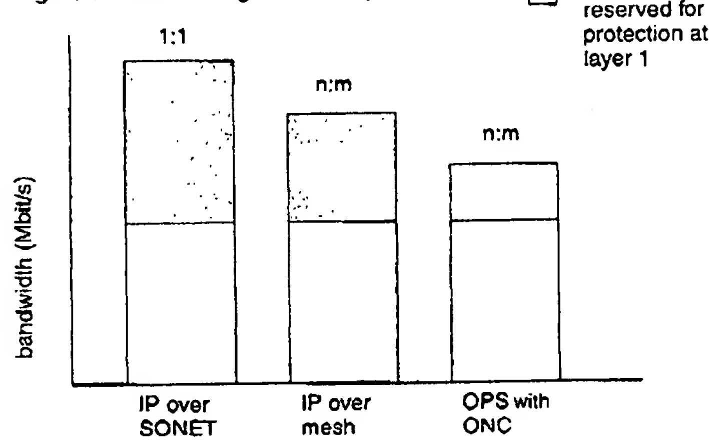
Figure 4 below illustrates the ONC's ability to increase the volume of traffic that can be sent by using more of the available bandwidth that cannot be accessed without an ONC. This holds true until the majority of the links on the network become highly congested. At this point, IP traffic following the open shortest path first (OSPF) protocol becomes more efficient than ONC-established MPLS paths. Although it would be extremely rare for a network to operate at such a congested level, should this occur, the ONC will prioritize shortest path traffic first (see Figure 4).

Figure 4: ONC vs. OSPF throughput



The ONC also achieves efficiency through the reduction of bandwidth required to be reserved for protection. The ONC has the ability to reserve resources for protection using both layer 1 and layer 3 resources, thereby reducing the amount of layer 1 resources needed to be reserved. Figure 5 compares the amount of protection required in an ONC-enabled network with that required in other network protection schemes.

Figure 5: ONC vs. ring and mesh protection



This chapter outlines how the ONC performs the following functions:

- multiprotocol label switching (MPLS) resource management
- adding, removing, and re-routing MPLS paths
- changing layer 1 connectivity

4.1 MPLS Resource Management

This section describes how the ONC manages MPLS resources. Specifically, it describes how the ONC adjusts the bandwidth of an MPLS path, and how the ONC balances traffic on two or more MPLS paths.

The ONC accesses the following inputs about the state of the network by polling the MPLS management information bases (MIB) located on the layer 3 core router, the OPTera Packet Core.

- MPLS path bandwidth - the total bandwidth of a given path
- MPLS path occupancy - the percentage of the total bandwidth of a given path that is being used.
- Router port link bandwidth - the total unallocated bandwidth on a link between two nodes in the network. A link can contain several MPLS paths.

The ONC also accesses the following data on the policy server:

- High-occupancy threshold - the maximum percentage of bandwidth in use on a given MPLS path before the ONC needs to take action to redistribute traffic.
- Low-occupancy threshold - the minimum percentage of bandwidth in use on a given MPLS path before the ONC needs to take action to redistribute traffic.

The ONC consults the above inputs and data at fixed intervals. In response to path occupancy levels that exceed the threshold values, the ONC performs two actions:

- First, the ONC adjusts the bandwidth of an MPLS path that has exceeded an occupancy threshold. If the path occupancy is too high, the ONC will increase the bandwidth on that path, provided that bandwidth is available on that link. If the path occupancy is too low, the ONC will decrease the

bandwidth on that path.

- Second, the ONC redistributes the traffic among the MPLS paths to ensure that the traffic is balanced relative to the amount of bandwidth on each path.

Bandwidth calculation

To adjust the bandwidth of an MPLS path that has exceeded an occupancy threshold, the ONC performs the following calculations:

for high occupancy threshold

To determine the increase in bandwidth that a path requires, the ONC makes this calculation: path occupancy minus path high occupancy threshold = percentage of bandwidth increase needed on that path.

For example, where the occupancy on a path is 90% and the high occupancy threshold is 80%, the bandwidth required on this path is 90 – 80 = 10%.

The ONC communicates this requirement for a 10 percent increase in bandwidth to the packet layer using the C9 interface to the OPTera Packet Switch (see "ONC to OPTera Packet Switch Interface: C9" on page 31). Specifically, the ONC uses Constraint-based Routed Label Distribution Protocol (CR-LDP). Using this protocol, the ONC sends a label request message containing a traffic parameter type-length-value (TLV), and a label-switched path identifier (LSPID) TLV. The combination of these two TLVs enable the ONC to increase the bandwidth on the designated path.

for low occupancy threshold

To determine the decrease in bandwidth that a path requires, the ONC makes this calculation: Path occupancy minus path low occupancy threshold divided by 4.

For example, where the occupancy on a path is 30% and the low occupancy threshold is 40%, the bandwidth required on this path is [30 – 40] /4 = -2.5% This calculation uses a denominator of 4 as a dampening factor to prevent a sharp decrease in bandwidth, in case traffic flow increases before the ONC can make another adjustment in bandwidth.

The ONC communicates this requirement for a 2.5 percent decrease in bandwidth to the packet layer using the C9 interface to the OPTera Packet Switch (see "ONC to OPTera Packet Switch Interface: C9" on page 31). Specifically, the ONC uses CR-LDP to send a label request message containing a traffic parameter type-length-value (TLV), and a label-switched path identifier (LSPID)

TLV. The combination of these two TLVs enable the ONC to decrease the bandwidth on the designated path.

Bandwidth share calculation

Once the ONC has adjusted the bandwidth on an MPLS path, it then redistributes the traffic among the MPLS paths with the same destination. To illustrate how the ONC balances the traffic, Figure 6 and Table 1 show four MPLS paths in an example network. Paths 1, 2, and 3 start at Node A and finish at node B. Path 4 starts at node A and finishes at Node C. The diagram shows the different routes the paths take to reach their destination.

Figure 6: MPLS Paths in Example Network

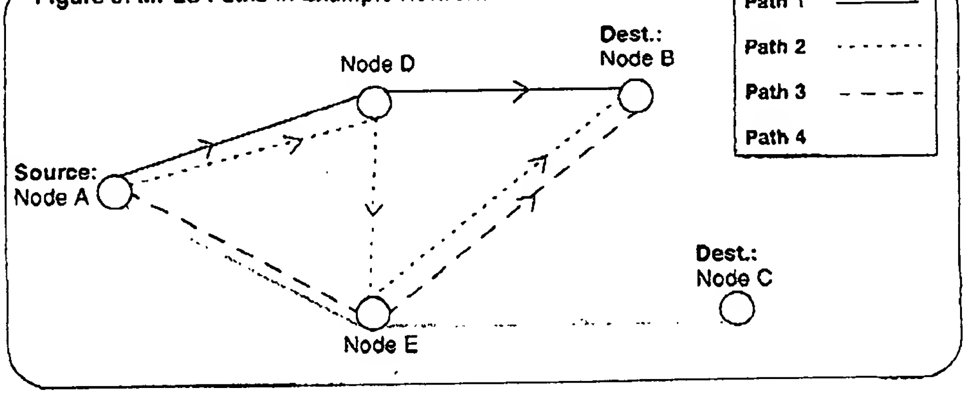


Table 1 shows how the ONC balances traffic among the MPLS paths shown in Figure 6. The paths in this network have a high occupancy threshold of 80% and a low occupancy threshold of 50%. The links are limited to 100 Mbit/s. Network changes are indicated in bold. ONC changes are indicated in the shaded cells.

Table 1: MPLS Resource Management

Time	Source	Dest.	Traffic-Node A to Node B (Mbit/s)	Traffic-Node A to Node C (Mbit/s)	Path	Links	Path Band-width (Mbit/s)	Share of traffic on path (%)	Actual traffic on path (Mbit/s)	Path occu-pancy (%)
7:00:00	Node A	Node B	50		1	A-D, D-B	20	22	11	56
	Node A	Node B			2	A-D, D-E, E-B	40	44	22.2	56
	Node A	Node B			3	A-E, E-B	30	33	16.7	56
	Node A	Node C		45	4	A-E, E-C	70	100	45	64
7:01:00	Node A	Node B	85		1	A-D, D-B	20	22	18.9	94
	Node A	Node B			2	A-D, D-E, E-B	40	44	37.8	94
	Node A	Node B			3	A-E, E-B	30	33	28.3	94
	Node A	Node C		45	4	A-E, E-C	70	100	45.0	64
7:01:10	Node A	Node B	85		1	A-D, D-B	28.8	25	21	73
	Node A	Node B			2	A-D, D-E, E-B	57.6	49	42.1	73
	Node A	Node B			3	A-E, E-B	30	26	21.9	73
	Node A	Node C		45	4	A-E, E-C	70	100	45	64
7:02:00	Node A	Node B	40		1	A-D, D-B	28.8	25	9.9	34
	Node A	Node B			2	A-D, D-E, E-B	57.6	49	19.8	34
	Node A	Node B			3	A-E, E-B	30	26	10.3	34

Time	Source	Dest.	Traffic-Node A to Node B (Mbit/s)	Traffic-Node A to Node C (Mbit/s)	Path	Links	Path Band-width (Mbit/s)	Share of traffic on path (%)	Actual traffic on path (Mbit/s)	Path occu-pancy (%)
	Node A	Node C		52	4	A-E, E-C	70	100	52	74
7:02:10	Node A	Node B	40		1	A-D, D-B	25.3	25	9.9	39
	Node A	Node B			2	A-D, D-E, E-B	50.5	49	19.8	39
	Node A	Node B			3	A-E, E-B	26.3	26	10.3	39
	Node A	Node C		52	4	A-E, E-C	70	100	52	74
7:03:00	Node A	Node B	50		1	A-D, D-B	20	22	11.1	56
	Node A	Node B			2	A-D, D-E, E-B	40	44	22.2	56
	Node A	Node B			3	A-E, E-B	30	33	16.7	56
	Node A	Node C		52	4	A-E, E-C	70	100	52	74

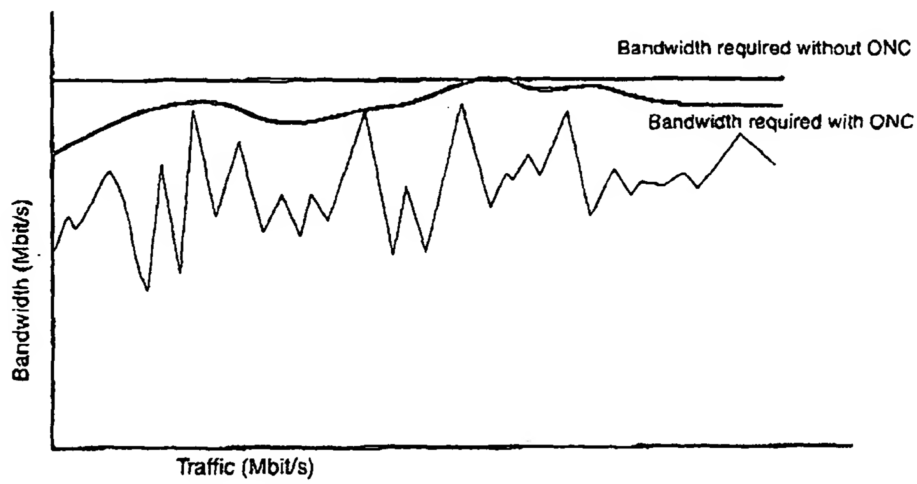
The calculation for rebalancing the traffic on existing MPLS paths is as follows:

path n bandwidth divided by the sum of the bandwidth of all the paths carrying traffic from the same source to the same destination.

For example, where the bandwidth of path 1 in Table 1 is 20 Mbits, divided by the sum of the bandwidth of paths 1, 2 and 3: 20+40+30= 20/90=22%. The total traffic on path 1 at time interval zero is 50 Mbits. 22% of 50 Mbits is 11 Mbits of traffic to be carried by path 1.

Figure 7 shows the traffic patterns on an MPLS path. Without the ONC, bandwidth is wasted—it must be kept at a constant high level to accommodate the peaks in traffic flow. When the ONC is in use, however, bandwidth is used efficiently and is kept to the minimum level required. Any excess bandwidth on the path can be made available to another path.

Figure 7: ONC Bandwidth Control



4.2 Adding, Removing and Re-routing MPLS Paths

This section describes how the ONC adds, removes or re-routes MPLS paths.

TO DO

4.3 Changing Layer 1 Connectivity

This section describes how the ONC accesses additional layer 1 resources when required by the packet layer.

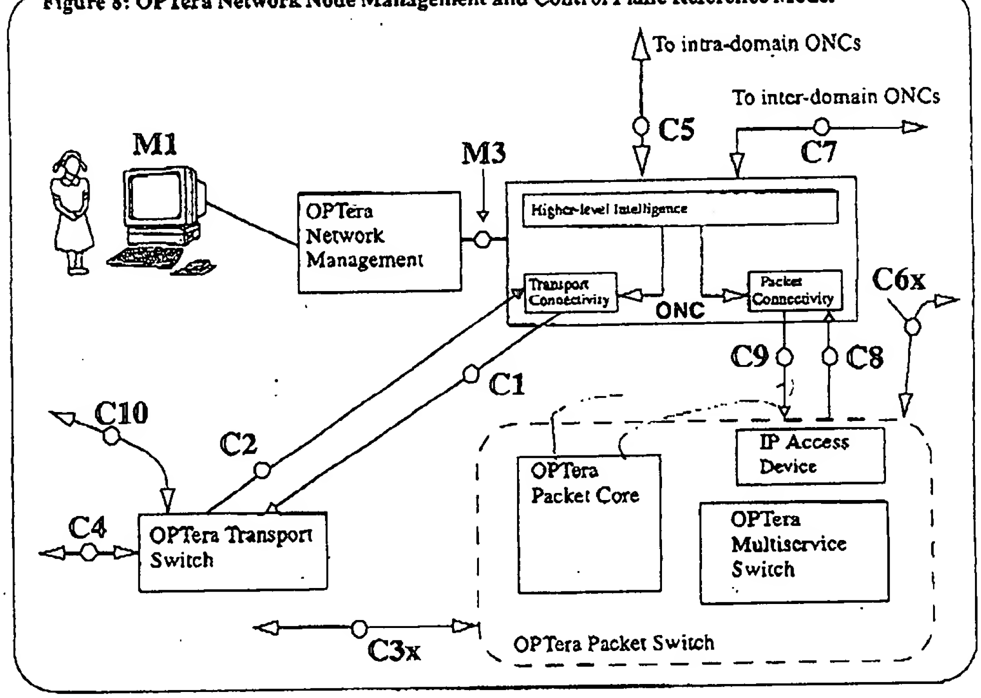
TO DO

6.1 Interfaces

This section describes the management interface and control interfaces used by the OPTera Network Controller. Figure 8 illustrates the management interfaces and control interfaces for the OPTera Packet Solution. This document focuses on the following interfaces:

- ONC-ONM interfaces: the means by which the ONC communicates with OPTera Network Management (see M3 in Figure 8)
- ONC-OPS interfaces: the means by which the ONC communicates with other components of the OPTera Packet Solution (see C1, C2, C8, C9 in Figure 8 below)
- ONC-ONC interfaces: the means by which the ONC communicates with other ONCs within its domain and with ONCs in other domains (see C5 and C7 in Figure 8)

Figure 8: OPTera Network Node Management and Control Plane Reference Model



6.1.1 Management Interface: M3

This is the interface between the ONC and the OPTera Network Management. Two management interfaces are provided:

- a simple command-line (text-based) interface with the syntax shown in "Command-Line Interface" on page 50.
- a textual computer-computer command interface that acts as a compressed encoding of the command-line interface. This format is described in detail in "Command Encoding" on page 51.

6.1.2 ONC to OPTera Transport Switch Interface: C1

This is a proxy-signaling interface from the ONC to the Automatically Switched Optical Network (ASON) for the establishment of source-routed connections. We assume an interface similar to a Constraint-based Routed Label Distribution Protocol (CR-LDP) interface to define the functionality required. From the description in reference (3), the subset of CR-LDP features required is as follows:

Label Request Messages including

- — Explicit Route (ER) Type-Length-Value (TLV)
- — Explicit Route Hop TLV
- — Traffic Parameter TLV
- — Preemption TLV

- Label Mapping messages including
  - — Traffic Parameter TLV

6.1.2.1 Explicit Route TLV

The Explicit Route TLV specifies the route of the optical path. The content of this TLV is a set of Explicit Route Hop TLVs. Given a maximum of 64 nodes per domain, an ONC would, at the most, use up to 64 Explicit Route Hop TLVs for one path.

6.1.2.2 Explicit Route Hop TLV

The assumption is that hops between the source and destination would be loosely defined with the OPTera Transport Switch's address instead of specific link and channel. The destination should be identified by its OPTera Transport Switch's address, link and channel (if required) identifier.

6.1.2.3 Traffic Parameter TLV

This TLV defines the path resources and its traffic capabilities.

For optical paths, the parameters required are path bandwidth and maximum delay.

6.1.2.4 Preemption TLV

This TLV specifies the preemption level of a given path. This can be used to identify paths for restoration that can carry preemptable traffic.

6.1.3 OPTera Transport Switch to ONC Interface: C2

This is a polled Operations, Administration, and Maintenance (OAM) interface that makes the following information available to the ONC:

- status information regarding connections established by ONC.

6.1.4 Intra-domain Interface: C5

This is the interface from the ONC to other ONCs within the same domain. It includes:

- the internal peer interface: C5.1 (TO DO)
- the internal hierarchical interface: C5.2 (TO DO)

6.1.5 Inter-Domain Interface: C7

This is the interface from the ONC to ONCs in other domains. This interface is not offered in the first phase of the product and is not defined in this issue.

6.1.6 ONC to OPTera Packet Switch Interface: C9

This interface is used by the ONC to control the OPTera Packet Core's Unified Label-Switched Paths. The interface uses SNMP to send CR-LDP requests to the OPTera Packet Core to:

- create new Unified Label-Switched Paths
- delete Unified Label-Switched Paths
- change traffic parameters of Unified Label-Switched Paths
- change the route of Unified Label-Switched Paths

From the description in reference (3), the subset of CR-LDP features required is

- Label Request Messages including
  - Explicit Route TLV
  - Explicit Route Hop TLV
  - Traffic Parameter TLV
  - Label-Switched Path Identifier (LSPID) TLV
  - Resource Class TLV
  - CR-LDP FEC TLV
- Label Mapping messages including
  - Traffic Parameter TLV

In the first release of the ONC, the C9 interface will also support resource reservation protocol (RSVP), a protocol that allows channels or paths in a network to be reserved for transmission of high-bandwidth messages. Supporting RSVP in the first release of the ONC is necessary to ensure compatibility with the first release of the OPTera Packet Core.

6.1.6.1 Explicit Route TLV

The Explicit Route TLV specifies the path of the LSP. The content of this TLV is a set of Explicit Route Hop TLVs. Given a maximum of 64 nodes per domain, an ONC would use, at the most, up to 64 Explicit Route Hop TLVs for one LSP.

6.1.6.2 Explicit Route Hop TLV

The ONC uses strict ER Hop specification. ER-Hop types required are IPv4 and LSPID.

6.1.6.3 Traffic Parameter TLV

This TLV defines the Unified Label-Switched Path resources and its traffic capabilities—the Rate, Peak Burst Size, Committed Data Rate, Committed Burst Size, Excess Burst Size—to define the bandwidth requirements of the Unified Label-Switched Path.

The negotiation flags allow downstream OPTera Packet Cores to negotiate (downwards) the resources allocation. This capability may be used by the ONC to get the largest path possible via a specific route.

The Traffic Parameter TLV is required for the Label Request message and the Label Mapping message (when negotiation has been requested).

TODO: if the packet switch does not allocate resources based on these parameters, then the ONC has to manage the "booking" at every hop. Is this done by passing the parameters via CR-LDP to the next hop that would check with its ONC?

TODO: Can the weight parameter be used to give the LSP a weighted-value for the route selection hashing algorithm where multiple LSPs are available for a given QoS and destination?

6.1.6.4 LSPID TLV

This TLV allows the ONC to modify the bandwidth of existing LSPs or their weighted value using the Action Indicator Flag of the LSPID TLV along with the Traffic parameter TLV.

TODO: We may use it also for rerouting LSPs?

6.1.6.5 Resource Class TLV

This TLV specifies which links are acceptable to this LSP. The 32-bit resource mask allows selection of any of 32 classes of links to be selected for the LSP.

This would allow the ONC to specify the type of protection of a physical (layer 1) link, such as a 1+1 protected versus non-protected link.

TODO: does this align with ASON Service Level Agreement definitions?

6.1.6.6 CR-LDP FEC TLV

TODO: What about the other FECs? We probably use exact HOST ADDRESSEs for Unified Label-Switched Paths.

6.1.7 OPTera Packet Switch to ONC Interface: C8

This is a polled OAM interface that makes the following information available to the ONC:

- State of Unified Label-Switched Paths
- State of ports (queues)
- State of links

The assumption here is that a link is the physical path that has been established between two OPTera Packet Cores. A port (queue) provides a specific service-

level (QoS) for a given link. Unified Label-Switched Paths provide a path to a specific egress OPTera Packet Core in the OPS network for a given service-level. In summary, for a given link, there are a number of queues (for QoS), and for each queue, there are a number of Unified Label-Switched Paths (for destination).

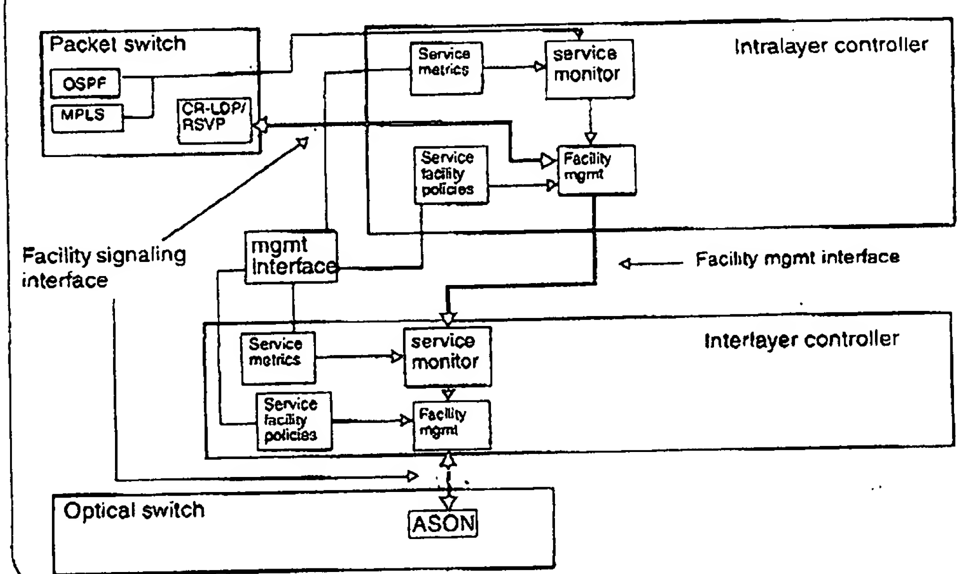
6.1.8 Flexible Service-Level Agreement Interface

This interface will allow the ONC to take advantage of flexible service level agreements at the packet level. This interface is not defined in this issue of the document.

6.2 Components

Each ONC has the structure shown in Figure 9; each of the components of that figure are described in the sub-sections below.

Figure 9: ONC components: high level structure



6.2.1 Management Interface

This provides an abstraction of the ONC for provisioning and other management purposes. It provides an interface to allow a higher-level management system or a local human operator to:

- provision the ONC:
  - — manipulate policies
  - — define the limits of the ONC's actions (e.g. report recommendations only, action recommendations automatically)
  - — set the ONC's IP address (which can only be performed through the local Command-Line Interface)
  - — set the nodes comprising the ONC domain and specify their network addresses
- access network statistics kept by the ONC:
  - — details of recommendations made with reasons
  - — current state of internal variables
  - — details of services that could not be carried because of lack of equipment (which could be used by the carrier for longer-term network planning)
- access ONC performance statistics
  - — messages exchanged between ONCs
  - — computation time for the various algorithms
  - — numbers of software faults
- receive alarms from the ONC when components fail
- receive information about attempted breaches of security (e.g. a program that tries to hack into and simulate an ONC component)

6.2.2 Service metrics

These are the parameters used to interpret and evaluate the health of a service node.

6.2.3 Service monitor

The service monitor receives information about the status of the service, and, based on the service metrics provided through the management interface, determines if action is required.

Intralayer role

The monitor receives statistics regarding the usage level of resources (unified packet paths) and makes changes to the topology of its network layer.

Interlayer role

The monitor receives statistics regarding usage level of the upper layer (that is, the link bandwidth allocated to unified packet paths). Between the packet and optical layers, for example, the service monitor receives information such as the packet layer's need for more optical resources.

6.2.4 Facility management

The facility management decides what action will be taken in response to the information received from the service monitor.

Intralayer role

The facility management determines:

- — how much bandwidth to add or delete for unified paths, and signals the packet switch to make the change
- — how to re-route an existing path using the topology information collected from the packet switch by the service monitor, and signals the packet switch to change the path
- — how to route a new path and what resource levels to allocate, and signals the packet switch to add the path
- — signals the packet switch to delete unused packet paths
- — provides statistics to the interlayer service monitor

Interlayer role

The facility management determines the parameters for new optical paths and signals the optical switch to establish the path. It also signals the optical switch to remove optical paths.

6.2.5 Facility Signaling Interface

The facility management uses this interface to signal its requests for actions.

Intralayer role

This is a command line interface (CLI) or simple network management protocol (SNMP) based interface:

— to the packet switch to request unified paths, change parameters, or delete paths. The parameters are based on either CR-LDP or RSVP

— to collect path statistics

— to collect topology information

— to collect link statistics

Interlayer role

This is an ASON-based interface to:

— request creation or deletion of optical paths

— receive connection status information

6.2.6 Facility Management Interface

This is an internal interface between the service monitor and the facility monitor at both the intralayer and interlayer levels of the ONC. This interface also permits the intralayer facility management to pass statistics to the interlayer service monitor.

6.2.7 Service facility policies

These are rules that govern the actions of the facility management.

Intralayer role

The policies define unified path profiles to meet service level agreements. These profiles are used when paths are allocated or modified. The policies also contain a list of unified path exceptions (those paths that should not be modified by the ONC as determined by the network administrator) used to filter the ONC's actions.

Interlayer role

The policies define link profiles and their relationship to unified path profiles. These profiles are used when optical connections are required. Again, the policies include a list of optical port exceptions to identify those ports on the packet switch that cannot be reconfigured at the optical layer.

6.2.8 Peer Interface

TBD

6.2.9 Inter-domain Interface

This can be thought of as an exterior gateway protocol. The interaction of ONCs between domains is limited and initially unsupported.

6.3 ONC Initialization

This section describes the data required for initialization of each nodal ONC. Each nodal ONC requires the pre-configured information listed in Table 2.

Table 2: Nodal ONC Configuration Information

Item	Learned From	Notes
ONC's own IP address	Pre-configuration during installation	From carrier's private IP address space.
OPTera Network Node identity	Pre-configuration during installation	Must be unique within the network.
Name Service address	Well-known	
Address of Management System	Nameserver using a well-known name	
Address of attached OPTera Packet Core	Nameserver using a well-known name	May not exist. From carrier's private IP address space.
Address of attached OPTera Transport Switch	Nameserver using a well-known name. Note that when the OPTera Transport Switch is not Nortel equipment the name and associated address may have to be manually configured in the nameserver.	May not exist. From carrier's private IP address space. This is the address that allows the ONC to access the OPTera Transport Switch's MIB (i.e. the SNMP address of the OPTera Transport Switch).
Capability of attached OPTera Transport Switch	Policy set manually	These include sufficient information for the nodal ONC to be able to create and initialize the appropriate objects.
Capability of attached OPTera Packet Core	Policy set manually	

6.3.1 Nodal ONC Initialization

TO DO.

This chapter attempts to scope the size of links and internal data structures (tables) required by the ONC. At this stage the estimates are approximate. They will be more precise in later versions of this document.

7.1 Inter-Node ONC Bandwidth

Bandwidth is required between ONCs for the exchange of information. This section attempts to estimate the required bandwidth.

The assumptions made for exchanges at the nodal-ONC layer once basic topology has been established are shown in Table 3.

Table 3: Bandwidth Sizing Assumptions

Item	Value	Notes
Nodes in area	64	
Unified Label-Switched Paths from my node	2048	64 nodes x 8 QoS x 8 routes
Length of statistics associated with a Unified Label-Switched Path	8 bytes	
Number of OPTera Transport Switch connections	256	
Length of statistics associated with OPTera Transport Switch connection	4 bytes	Limited to status
Information exchange frequency	10 Hz	This is probably too high for initial implementations

The required bandwidth with these assumptions would be approximately

$$10 \times ((2048 \times 8 \times 8) + (256 \times 8 \times 8)) = (1.5) \times 10^6$$

(i.e. 1.5Mbit/s).

Note that this assumes only one area for the ONC domain.

7.2 Table Sizing

TO DO

7.3 Engineering parameters

This section defines the variables such as fixed intervals, thresholds, and engineered parameters of the algorithms.

TO DO

This appendix describes a hierarchical model for the arrangement of ONCs in a large network. In the initial release of the ONC in a 64-node network, however, this hierachical model will not be necessary.

B.1 Terminology

This section defines the terms domain, area and node.

A **domain** is a carrier's total administrative zone. For traffic engineering reasons, a carrier divides its domain into a number of **areas**. Each area consists of a collection of **nodes**. In the context of this document, a node refers to an OPTera Network Node, which contains one or more of the following: OPTera Packet Switch, OPTera Network Controller, OPTera Transport Switch, OPTera Multiservice Switch, and an OPTera IP Access Device.

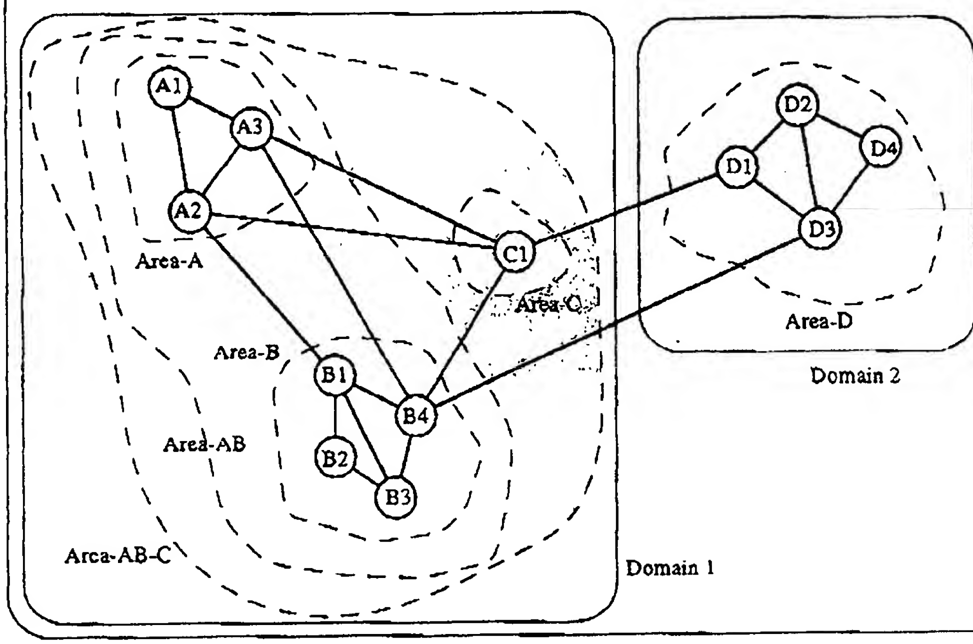
The interaction of ONCs between domains is limited and initially unsupported. This inter-domain interface can be thought of as an exterior gateway protocol.

Within an area the optimization required of the ONC is greater than it is between areas.

In a large network, areas may themselves be grouped hierarchically. The number of levels of hierarchy in any domain must be the same for all nodes. The highest level in the hierarchy corresponds to the domain itself. As a minimum, a domain must contain at least one area.

Consider the network shown in Figure 10. Domain 1 includes five areas: A, B, C, AB, and AB-C. Area A is comprised of three nodes: A1, A2, A3. Area B is comprised of 4 nodes: B1, B2, B3, B4. Area C is comprised of one node: C1. Area AB is the combination of Area A and Area B. Area AB-C is the combination of Areas A, B, and C. Domain 2 comprises one area: D.

Figure 10: Domains and Areas: Physical View



B.2 ONC Layering

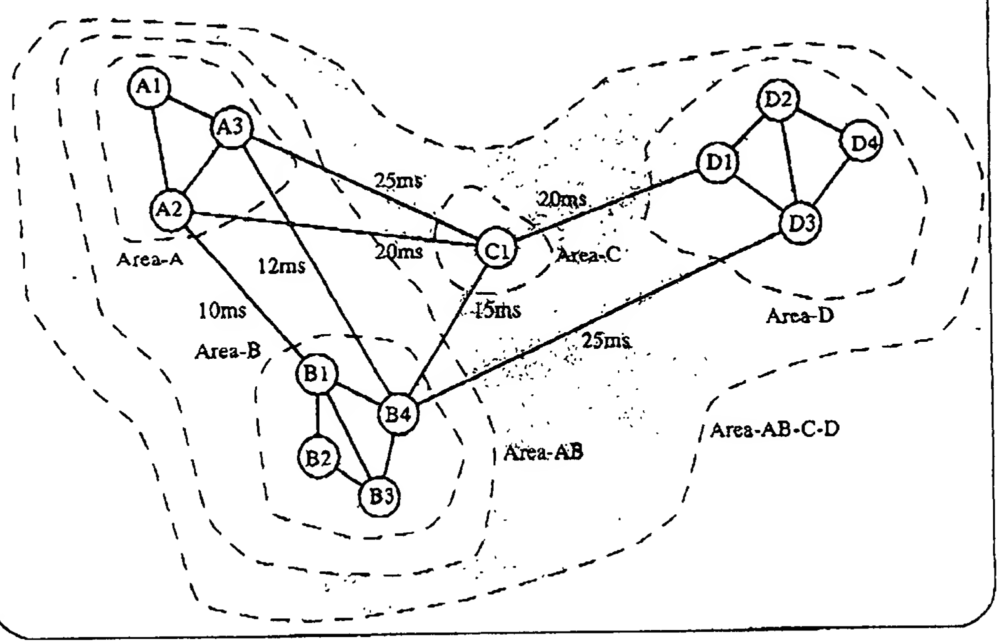
Two independent dimensions of layering are present in the ONC:

- the ONC controls several layers of the telecommunications network: fibre, wavelength, label-switched path, etc.
- the ONC itself is layered into hierarchical areas.

This section deals only with the second of these layerings.

Consider the network shown in Figure 11. This represents a geographically-distributed network of 12 OPTera Network Nodes. The nodes have been grouped into areas in such a way that the boundaries make organizational sense to the carrier and logical sense as regards grouping by delay.

Figure 11: Example Network



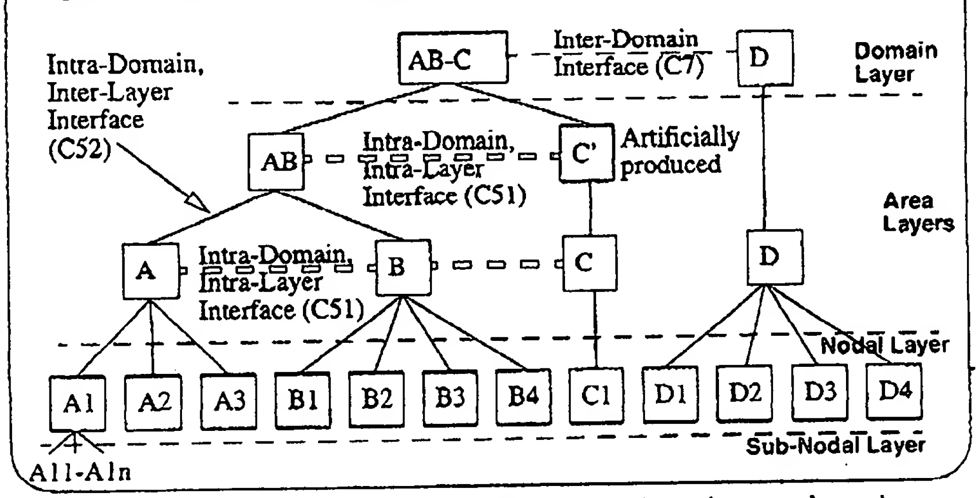
The nodes within Area-A are connected by links with latency that does not affect the ONC's algorithms (below about 1ms or 300km). The nodes within Area-B and Area-D are also so connected. Area-AB is a combination of Area-A and Area-B. Area-AB has been designed to conform to the carrier's organizational structure, as has Area-AB-C-D, which is a combination of Area-AB, Area-C and Area-D.

At least one ONC is associated with each node. In addition, there is an ONC associated with each area, arranged in a hierarchy as illustrated in Figure 12.

The roles of each of the ONC layers are as follows:

- sub-nodal ONCs: these are responsible for the collection of detailed information at a low-level (typically associated with a link or Unified Label-Switched Path). They perform calculations to characterize the state of the low-level connections and pass this information to the nodal-level ONCs. In general, sub-nodal ONCs do not perform network rebalancing. Rules at this level are likely to be common (independent of network or node topology), remain stable and membership functions (the value of a

Figure 12: ONCs in Example Network: Logical Diagram



parameter for a certain variable) are likely to change only rarely.

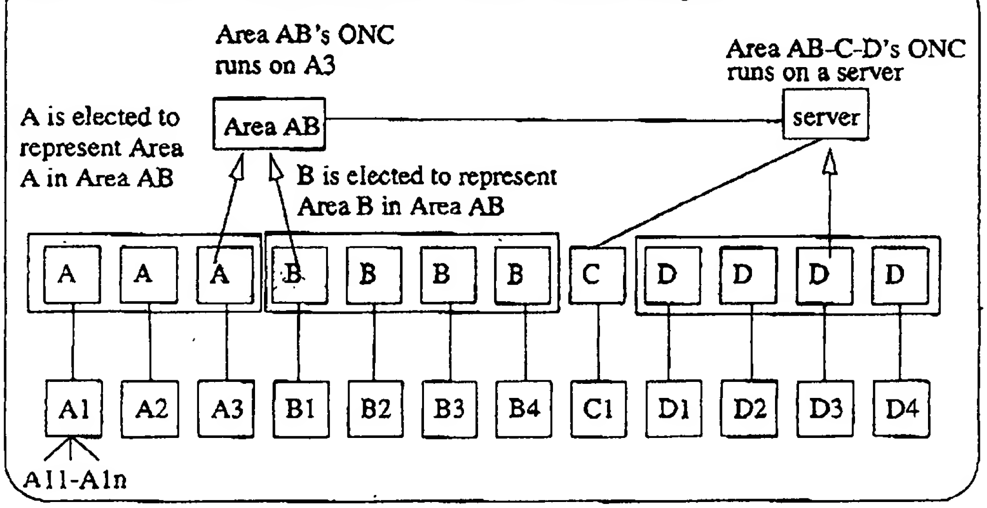
- nodal ONCs: these are responsible for analyzing the information received from the sub-nodal ONCs and making node-level decisions. A node-level decision involves rebalancing without modifying network topology. In general, rules at this level are largely common (network-independent), are stable and membership functions are likely to change frequently to maintain discrimination—the effect of parameter changes on the rules—as conditions change.
- area ONCs: there are any number of these in the hierarchy. They are responsible for making area-wide decisions that could not be handled at the nodal level. Rules at this level are changed frequently to reflect network topology changes (as they embed the network topology) and membership functions are likely to change frequently to maintain discrimination.

Thus:

- the decision to route traffic from a failed link to another, spare, link on the same card might be taken at the sub-nodal level.
- the decision to adjust the fractions of traffic using two different links out of a node might be taken at the nodal level.
- the decision to route traffic through A1-A2-B1-B4 to offload A1-A3-B4 would require agreement at the Area-AB level without reference to Area-AB-C-D.

In summary, each ONC is responsible for summarizing the information at its level, making rebalancing decisions at its level and providing its parent with the summarized information.

Figure 13: ONCs in Example Network: Possible Physical Diagram



Distribution of the code is, to some extent, a design decision and it is an architectural principle that it be flexible, but it is expected that the ONCs would be distributed as follows:

- sub-nodal ONCs: these reside as close to their interface points as possible, preferably on the linecards themselves and, if that is not possible, then on the nodal processor
- nodal-ONCs: these run on the corresponding nodal processors
- area-ONCs: these run in one of three ways:
  - — on all of the nodes within their area. This is likely to be true of at least the first few layers of area hierarchy. The algorithms designed to allow such distributed ONCs to run are as yet undefined.
  - — on one of the nodes in the area. In this mode one ONC in each of the lower-layer groups is nominated (in PNNI-fashion) to act as the group's representative at the higher level. The algorithm for choosing the representative and reelecting the representative on failure is as yet undefined. This technique may be inappropriate in a geographically small network (where the number of layers may not warrant it) or in a global network (where the delays may prohibit it) but it may be required in continental-scale networks.
  - — on a centralized server that is not part of a node. In this mode, as in the previous mode described above, one ONC in each of the lower-layer

groups is nominated to act as the group's representative at the higher level.

This technique is illustrated in Figure 13 where the network of Figure 12 has been mapped to a possible physical instantiation. In Figure 13, it has been decided that Area-A, Area-B, Area-C, and Area-D should be run in distributed mode, but that Area-AB and Area-AB-C-D should be run in centralized mode. Thus, for example, the ONCs for sub-nodes A11, A12, ..., A1n, A1, and A all run on the processor associated with a single node, but there is only one ONC associated with Area-AB-C-D, which runs on a server in a centralized location. Area AB's ONC runs on A3.

ONC interfaces do not occur at all layers (see Table 4).

Table 4: ONC Interfaces

Interface	C1	C2	C5.1	C5.2	C7	C8	C9	M3
Sub-Nodal		✔		✔		✔		✔
Nodal	✔		✔	✔			✔	✔
Area			✔	✔				✔
Top-level area (domain)				✔	✔			✔

BACKGROUND OF THE INVENTION

Field of the Invention

The invention is directed to communication networks, and in particular to a controller for an optical communication network.

Background Art

The explosive growth of data networks, particularly the Internet, presents both tremendous opportunities and challenges for service providers. Today's networks are struggling to keep up with the demand created by new users, new technologies and new high-bandwidth applications. To meet this ever-expanding market demand, service providers are re-evaluating how to configure their networks. Traditional networks, developed using an overlay model where layers are built and managed independently, make it difficult for service providers to optimize network resources and provide reliable networks that are also cost-effective.

In addition, the growth in the demand for bandwidth and the shifting nature of traffic patterns create a need for networks that are flexible and scalable in size, cost-effectiveness and able to provide for efficient management of network resources. Currently, to address bandwidth requirements that cause variability in network traffic characteristics, service providers have little choice but to engineer their networks for "worst-case" traffic volumes, resulting in under-utilized network resources. When changing traffic patterns require reconfiguring their networks, service providers must manually engineer and provision new connections at both the packet and optical layers of the network.
Use of SONET/SDH for automatic protection switching in case of physical link failures also requires that large amounts of spare bandwidth be held in for use in the event of a fault.

The complexity of the traditional network built using a multi-layer model also makes it more difficult for service providers to identify opportunities for network optimization. An IP network (layer-3), for

example, may run over a frame relay network, which runs over an ATM network (layer-2), which runs over a SONET/SDH network (layer-1), which in turn runs over an optical/wavelength network (layer-0). This multilayer model allows each layer to evolve independently, while continuing to support legacy services. However, the large number of different devices in the network complicates the development, deployment and management of new services.

Furthermore, each layer of the network typically has an independent management structure and associated processes that only have visibility of the topology and state information of that particular layer. The number of management systems increases network cost, while adding complexity to the network-wide operation tasks such as provisioning, performance monitoring and fault isolation. Two primary difficulties with the multilayer model are the inability of service providers to optimize the use of network resources and to provide the network with cost-effective reliability. These difficulties are the driving forces behind the need for network intelligence.

SUMMARY OF THE INVENTION

Service providers typically have to dedicate teams of network architects to keep up with network growth and changes to the network. The ONC reduces operations costs through automation of traffic engineering and through simplified cross-layer management. The ONC gives service providers the ability to map their business logic into their network operations. Policy information is used by the ONC to make network decisions that can prioritize the highest revenue-generating traffic and services.

By dynamically optimizing network resources, the OPTera Network Controller allows service providers to manage services rather than networks. The result for service providers is increased revenues, decreased capital costs, and new opportunities for revenue generation.

The ONC gives service providers the ability to map their business logic into their network operations. Policy information is used by the ONC to make network decisions that can prioritize the highest revenue-generating traffic and services. Service providers typically have to dedicate teams of network architects to keep up with network growth and changes to the network. The ONC reduces the operation costs through automation of traffic engineering and through simplified cross-management.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, features and advantages of the invention will be apparent from the following more particular description of the preferred embodiments, as illustrated in the appended drawings, where:

Figure 14 is an example of a network provided with the ONC according to the invention;

Figure 15 shows a high level structure, illustrating the ONC components;

Figure 16A is a bandwidth-traffic graph for a typical network;

Figure 16B is a bandwidth-traffic graph illustrating the ONC control;

Figure 17 shows an example of how the ONC balances the traffic between four MPLS paths in an example network;

Figure 18A is a graph for comparing the ONC versus OSPF throughput;

Figure 18B is a graph showing the bandwidth usage for networks using ONC, versus ring and mesh protection;

Figure 19 is an example of the physical view of a network provided with ONC, for defining terms domain, area and node;

Figure 20A is another example of the physical view of a network provided with ONC;

Figure 20B shows the logical view of the example network of Figure 20A; and

Figure 21 is an example of the physical view of a current network.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Smart packet/optical interworking means the ability for the optical and packet layers of the network to share information, take action based on this information, and to optimize the network dynamically. The specification discloses a network that carries Internet over the optical layer, using multi-terabit switching and routing capacities, dynamic network optimization, consolidation of multiple protocols onto a unified network, and carrier-grade reliability.

A fundamental component of this network is the network controller, called herein OPTera Network Controller (ONC). ONC is a smart packet/optical inter-working system that uses topology information from the packet and optical layers and network wide policy information to react to changing network configuration adjustments.

In this specification, the term "domain" is used for a carrier's total administrative zone. For traffic engineering reasons, a carrier divides its domain into a number of 'areas'. Each area consists of a collection of nodes. In the context of this document, a 'node' refers to a network node which comprises one or more of the following: packet Switch, a nodal network controller, a transport switch, a multiservice switch, and an IP access device.

Figure 14 is a block diagram of a network provided with the ONC 10 according to the invention. Network 1 comprises in this example three nodes, but it is to be understood that ONC 10 may serve networks with a much larger number of nodes. Also, the configuration of a nodes is by way of example; what should be noted is the presence of a layer-3 core router and a layer-1 transport switch. The ONC 10 works at three levels to optimize network resources:

1. At the packet layer 1, the ONC 10 balances the traffic, adjusts the bandwidth of multiprotocol label-switching (MPLS) paths, and automatically sets up end-to-end paths.

2. At the optical layer **2**, the ONC **10** defragments the optical network, allowing the network to make use of previously stranded resources.

3. Between the optical and packet layers, the ONC **10** works to allow optical resources to be used directly by the packet layer to respond to congestion or increased demand at the packet layer.

Figure 14 also shows instances of the OPC at each node, i.e. OPCI 12-1, 12-2 and 12-3 respectively. The network manager provides integrated management of these components, using interfaces by which a nodal ONC within a domain communicates with nodal ONCs in the same domain or in other domains.

A service node provides a path-oriented service, is managed by a service level agreement, and engineers network resources to meet the agreement at the optimal cost. As shown in Figure 14, at layer-3, a service node comprises a packet core router **7-1**, **7-2** or **7-3**, and it may also comprise edge routers or multi-service switches **6-1**, **6-2**, **6-3** and **6-4**. The edge routers **6-1** and **6-3** act as a peripheral to provide IP interfaces to the respective node. The multiservice switches **6-2** and **6-4** act, for example, as a peripheral on the associate core router to provide ATM, IP, or Frame Relay interfaces.

The ONC **10** interacts with layer-3 equipment to gain a global view of the network through automatic discovery of network topology, and by accessing data regarding network traffic levels and status of the network. ONC also optimizes network traffic by changing the MPLS path characteristics by creating new MPLS paths, and by aggregating MPLS paths.

A facility node is a lower level resource required by the service node. The ONC allows the packet layer to use the optical resources in response to congestion or increased demand at the packet layer. At layer-1, the node comprises an optical switch **8-1**, **8-2** or **8-3**. The switches are for example programmable and capable of switching at least one fiber, one wavelength, wavelength band, or SONET/SDH frames. The network is also provided with Automatically Switched Optical Network

(ASON) functionality shown at 20. The term ASON is used for the protocols within the transport domain to determine topology and establish signalling connections.

As indicated above, the ONC functions can be divided into two categories, namely intra-layer functions and inter-layer functions.

Regarding inter-layer functionality, the ONC provides a network with intelligent dynamic resource management between a service node (i.e. a node of the layer-3 network) and a facility node (i.e. a node of the layer-1 network).

The interface **16** in the direction from the ONC **10** to the transport switch **8** could be for example a proxy signalling interface to the ASON **20** for the establishment of source-routed connections. This interface may be similar to a CR-LDP (constrain-based routed label distribution protocol) interface to define functionality required. The label request messages may for example include explicit route type-length-value (ER_TLV), explicit route hop TLV, traffic parameter TLV, and preemption TLV. The label mapping messages may for example include traffic parameter TLV.

ER-TLV specifies the route of the optical path. The content of this TLV is a set of explicit route hop TLV's. Given a maximum of 64 nodes per domain, an ONC would, at the most, use up to 64 explicit route hop TLV's for one path.

The assumption for explicit route hop TLV is that hops between source and destination would be loosely defined with the address of the transport switch instead of specific link and channel. The destination could be identified by the switch address, link and channel identifiers.

The traffic parameter TLV defines the path resources and its traffic capabilities. For optical path, the parameters required are path bandwidth and maximum delay.

The preemption TLV specifies the preemption level of a given path. This can be used to identify paths for restoration that can carry preemptable traffic.

The interface **16** in the direction from the transport switch **8** to the ONC **10** performs for example a pooling OAM (operation, administration

and maintenance) interface that provides the ONC with status information regarding connections established by ONC.

The ONC to core router interface 14 is used by the ONC to control the unified label-switched paths (ULSP). The interface uses SNMP (simple network management protocol) to send CR-LDP requests to the layer-3 service node to create or delete ULSP, change traffic parameters and/or routes of UPLS's. The label request messages include explicit route TLV, explicit route hop TLV, traffic parameter TLV, label-switched path identifier (LSPID) TLV, resource class TLV, CR-LDP FEC TLV. The label maping messages include traffic parameter TLV. This interface may also support resource reservation protocol (RSVP), a protocol that allows channel or paths in a network to be reserved for transmission of high-and messages.

The explicit route TLV specifies the path of the LSP. The content of this TLV is a set of explicit route hop TLV's. Given a maximum of 64 nodes per domain, an ONC would use, at the most, up to 64 explicit route hop TLV's for one LSP.

The ONC uses strict ER hop specification. ER-hop types required are IPv4 and LSPID.

Traffic parameter TLV defines the ULSP resources and its traffic capabilities-the rates, peak burst size, committed data rate, committed burst size, excess burst size, to define bandwidth requirements of the ULSP. The negotiation flags allow downstream the service nodes to negotiate (downwards0 the resources allocation. This capability may be used by the ONC to get the largest path possible via a specific route. However, if the service node does not allocate resources based on these parameters, then the ONC has to manage the booking at every hop.

Eventually, the weight parameter may be used to give the LSP a weighted value for the route selection hashing algorithm where multiple LSP's are available for a given QoS (quality of service) and destination.

The traffic parameter TLV is required for the label request message and the label mapping message, when negotiation has been requested.

The LSPID TLV allows the ONC to modify the bandwidth of existing LSPs or their weighted value using the action indicator flags of the LSPID TLV along with the traffic parameter TLV.

The resource class TLS specifies which links are acceptable to this LSP. the 32-bit resource mask allows selection of any of 32 classes of links to be selected for the LSP. This would allow the ONC to specify the type of protection at the physical layer (layer-1) link, such as 1+1 protection versus non-protected link.

CR-LDP FEC TLV preferably uses exact host addresses for ULSP.

The interface between the service node and the ONC is again a pooled interface that makes available to the ONC information about the state of the ULSPs, state of ports and state of links. The assumption is that a link is the physical path that has been established between the service nodes. A port (queue) provides a specific service-level (QoS) for a given link. ULSP provide a path to a specific egress service node for a given service-level. In summary, for a given link, there are a number of queues (for QoS) and for each queue, there are a number of ULSPs (for destination).

The ONC uses the ASON interface **20** to gain access to the optical transport layer **2** to map packet layer service requirements to available optical layer resources. This is done by automatically setting up the end-to-end physical path topology for the layer-3 network using configuration data such as traffic parameters provided by a network manager 4. The ONC also dynamically sets-up, tears-down or alters connections at the layer 1 level and de-fragments the optical network, thereby allowing the network to make use of previously stranded resources. As indicated above, the ONC also allows interaction between the optical and packet layers, by allowing optical resources to be used directly by the packet layer to respond to congestion or increased demand at the packet layer. Finally, the ONC uses the ASON interface **20** at different sub-layers within the optical layer, i.e. between an optical switch and a lambda switch by allowing the optical switch to use the resource of the lambda switch.

Network manager 4 provides a unified view of the network for the collection of devices in the network, and offers the human interface to the network information. Via management interface **19**, network operators can control the ONC by means of user-defined policies. The policies may be adjusted so that the network operator can customize the ONC's actions. The network operators can also obtain the records of audit trails and explanations of every action performed, and obtain recommendations from the ONC of potential actions to be performed when the ONC decides that human intervention is required. Interface **19** may for example include a simple command line interface, and a textual computer-computer command interface that acts as a compressed encoding of the command-line interface.

Although the block diagram illustrates an implementation of the ONC in the network, the interfaces **14-1**, **14-2** and **14-3** to the core routers and interfaces **16-1**, **16-2** and **16-3** to the optical transport equipment are designed to be open interfaces such that ONC can be used in an open environment.

Regarding the intra-layer functionality, at the service node layer, the ONC understands the service level agreements (SLA's) and manages resources at this layer. At the optical layer, the ONC defragments the optical network thereby allowing the network t make use of previously stranded resources. Also, at different sub-layers in the optical layer for example between an optical switch and a lambda switch, by allowing the optical switch to use the resources of the lambda switch. Intra-domain interfaces **15-1** and **15-2** may include for example an internal peer interface and an internal hierarchical interface.

The ONC is designed to run on any operating system and performs at least the following functions:

1. Automatically creates physical connectivity for the layer-3 network based on user- defined policies (such as quality of service) via the ASON interface.

2. Automatically creates MPLS connectivity between the layer-3 network based on user defined policies (such as quality of service) and user-defined traffic requirements.
3. Automatically optimizes packet traffic services based on user-defined policies by managing layer-3 and layer-1 resources.
4. Collects and reports network utilization information.
5. Reports recommendation for actions to network operators.
6. Executes MPLS path protection based on quality of service agreement defined by user.

The block diagram of the OPC is shown in Figure 15. OPC has two major functional blocks: maintenance and service.

The maintenance block **21** provides regular maintenance capability such as software upgrades or alarm reporting. The service block **22** is comprised of several sub-functional blocks; metrics database **23**, policy database **24**, filters **25**, control plug-ins **26**, metric monitoring **27**, optimization and control **28**, packet equipment interface **14** and ASON interface **16**.

As discussed above, management interface **19** provides an abstraction of the ONC for provisioning and other management system or a local human operator. The interface manipulates policies, defines the limits of the ONC's auctions (e.g. report recommendations only, action recommendations automatically), sets the ONC IP addresses through the local command line interface, and sets the nodes comprising the ONC domain and specify their network addresses. Access network statistics are kept by the ONC. This may for example include details of recommendations made with reasons, current state of internal variables, details of services that could not be carried because of lack of equipment.

The management interface **19** also accesses the ONC performance statistics, such as message exchanged between ONC's, computation time for the various algorithms, numbers of SW faults. In addition, IF **19** receives alarms from the ONC when components fail, receive information about attempted breaches of security.

Metrics database **22** and policy database **23** are controlled via network manager **3**. The metrics database stores parameters used to evaluate traffic conditions to determine if action should be taken by the ONC. Examples are threshold values and data sampling intervals.

Policy databases **23** store rules for the control algorithm. These rules are used by the ONC to decide if the action to be performed is appropriate.

Plug-ins **25** are used to enable more flexible control. They also allow algorithms to be upgraded without upgrading the software. Control algorithm plug-ins allow different types of controls to be applied to different scenarios.

Filters **24** are algorithms that provide additional data processing for collected data before or after the comparison to eliminate transient conditions.

Metrics monitoring **26** is a data collection and processing engine. It interacts with the packet equipment to query/collect necessary performance metrics, filters data if required and uses data from the metrics database for comparison. When necessary, results will be fed into the control functional block for action. Metrics monitoring **26** also reports some of the performance metrics to the network operator.

Optimization and control **27** performs all the optimization actions and resource management functions. All actions are governed by the policy database. The algorithm being used to interpret the database is provided by the plug-in. Different algorithm functions include building the MPLS mesh or building the layer 1 connectivity between routers. When an action is to be taken, it uses the packet equipment or the ASON interface to send out commands to the appropriate layer.

Packet equipment interface **14** and ASON interface **16** were discussed above in connection with Figure 14. These two blocks interface to the packet equipment and transport equipment. They perform functions such as protocol encoding and actual message sending. They provide actual interface decoupling functionality such that the ONC can adapt to

interface changes or interface with different equipment that does not support standard interfaces.

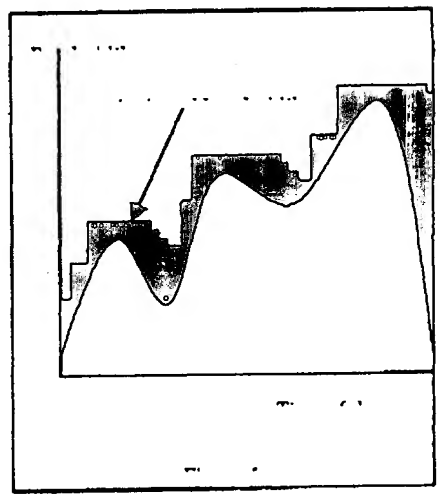
To maintain a certain level of service grade, the bandwidth between two nodes must be engineered to a level that will satisfy traffic requirements most of the time or all the time. In other words, a link is always over-engineered and therefore most of the time under-utilized. This is illustrated in Figure 16A, which shows a bandwidth-traffic graph for a typical network. The shaded portion under the engineered level represents the resources not being used. The total amount of bandwidth being over engineered is very significant in the network. If this bandwidth can be tapped, the amount of traffic being carried in the same network can be increased without sacrificing grade of service. This will translate to increased revenue for service providers.

ONC taps this bandwidth by using logical connections (MPLS paths) between two points and dynamically adjusting the bandwidth of an MPLS path according to the path utilization level (patent pending). This MPLS path bandwidth adjustment results in the following advantages over the current networks, as shown also in Figure 16B. Namely:

1. bandwidth assigned between two points is adjusted according to traffic level, not statically assigned and therefore conserves bandwidth for other connections.
2. packet loss will be reduced until the bandwidth demand exceeds a certain level, leading to a higher performance network
3. if packets are dropped due to congestion, they will be dropped at the entry point of an MPLS path and will not affect the intermediate nodes as is the case with the open shortest path first (OSPF) protocol. This behavior allows the ONC to monitor network performance by measuring the packet loss at the ingress point of a router.

Figure 16B is a bandwidth-traffic graph illustrating the ONC control. When the ONC detects traffic level rises beyond a configurable threshold (via the network manager), the ONC will increase the bandwidth assigned to that path. When the traffic level drops below a configurable threshold, the ONC decreases the bandwidth of the path, but by a value smaller than

the previous increment. The algorithm used will closely track traffic patterns and therefore maximize the bandwidth conservation effect. As a result, network performance is optimized by dynamically adjusting bandwidth supply to bandwidth demand.



Using dynamic bandwidth assignment to an MPLS path, the ONC effectively increases network resources and maximizes the network traffic by redistributing resources to/from other MPLS paths. It adjusts resources in three steps:

When a path's utilization exceeds the pre-configured threshold, the ONC will first attempt to increase the bandwidth of the path. The amount of bandwidth to be increased is based on a patent pending algorithm, such that the average utilization of each path between two nodes is maintained at the same level.

If there is no bandwidth left to increase the size of paths between two nodes, the ONC will attempt to create a new MPLS path using a different physical route.

If the first two attempts are not possible, the ONC will use ASON services to create a new optical route and thereby change the physical transport topology and create more resources.

The ONC accesses the following inputs about the state of the network by pooling the MPLS management information databases (MIB) located at the layer-3 service node, i.e core router 7. We define the 'MPLS bandwidth' as the total BW of a given path, 'MPLS path occupancy' as the percentage of the total BW of a given path that is being occupied, and 'the router port link bandwidth' as the total unallocated BW on a link between two nodes in the network.

'High occupancy threshold' in this specification is the maximum percentage of BW in use on a given MPLS path before the ONC needs to

take action to redistribute traffic. 'Low occupancy threshold 'is the minimum percentage of BW in use on a given MPLS path before the ONC needs to take action to redistribute traffic.

The ONC consults the above inputs and data at fixed intervals. In response to path occupancy levels that exceed the threshold value, the ONC performs two actions:

-First, the ONC adjusts the BW of an MPLS path that has exceeded an occupancy threshold. if the path occupancy is too high, the ONC will increase the BW on the path, provided that BW is available on that link. If the path occupancy is too low, the ONC will decrease the BW on that path.

-Second, the ONC redistributes the traffic among the MPLS paths to ensure that the traffic is balanced relative to the amount of BW on each path.

To adjust the BW on an MPLS path that has exceeded an occupancy threshold, the ONC performs the following calculations:

High occupancy threshold:

To determine the increase I (in percentage) in the bandwidth that a path requires, the ONC calculates:

$$PO - HOT = I, \qquad \text{EQ1}$$

where PO is path occupancy, HOT is the high occupancy threshold.

For example, where the occupancy on the path is 90% and the high occupancy threshold is 80%, the increase in bandwidth required for this path is 90%-80%=10%.

The ONC communicates this requirement for a 10% increase in bandwidth to the packet layer, using CR-LPD (constraint-based routed label distribution protocol). Using this protocol, the ONC sends a label request message containing a traffic parameter TLV (type-length-value), and a LSPID (label-switched path identifier)TLV. The combination of

these two TLVs enables the ONC to increase the bandwidth on the designated path.

Low occupancy threshold:

To determine the decrease D in bandwidth that a path requires, the ONC calculates:

$$(PO - LOT)/4 = D, \qquad \text{EQ2}$$

where LOT is the low occupancy threshold.

For example, where the occupancy on the path is 30% and the low occupancy threshold is 40%, the decrease in bandwidth required for this path is (40-30)/4=2.5%.

The ONC communicates this requirement for a 2.5% decrease in bandwidth to the packet layer, using CR-LPD to send a label request message containing a traffic parameter TLV, and a LSPID TLV. The combination of these two TLVs enables the ONC to decrease the bandwidth on the designated path.

Bandwidth share calculation

Once the ONC has adjusted the bandwidth on a MPLS path, it redistributes the traffic among the MPLS paths with the same destination. To illustrate how the ONC balances the traffic, Figure 17 and Table 1 show four MPLS paths in an example network. Paths P1, P2 and P3 start at node A and finish at node B. Path P4 starts at node A and finishes at node C. The diagram shows the different routes the paths take to reach their destination.

Table 1 shows how the ONC balances traffic among the MPLS paths shown in Figure 17. The paths in this network have a high occupancy threshold of 80% and a low occupancy threshold of 50%. The links are limited to 100Mbit/s. Network changes are indicated in bold. ONC changes are indicated in italic.

TABLE 1: MPLS resource management

Time	Source	Dest.	%BW A-B	%BW A-C	Path	Links	Path BW	Share BW/path	BW/path	PO
00:00	A	B	50		1	A-D D-B	20	22	11	56
	A	B			2	A-D D-E E-B	40	44	22.2	56
	A	B			3	A-E E-B	30	33	16.7	56
	A	C		45	4	A-E E-C	70	100	45	64
01:00	A	B	85		1	A-D D-B	20	22	18.9	94
	A	B			2	A-D D-E E-B	40	44	37.8	94
	A	B			3	A-E E-B	30	33	28.3	94
	A	C			4	A-E E-C	70	100	45	64
01:10	A	B	85		1	A-D D-B	*28. 8*	*25*	21	73
	A	B			2	A-D D-E E-B	*57. 6*	*49*	42.1	73
	A	B			3	A-E E-B	30	*26*	21.9	73
	A	C		45	4	A-E E-C	70	100	45	64
02:00	A	B	40		1	A-D D-B	28. 8	25	21	34
	A	B			2	A-D D-E E-B	57. 6	49	19.8	34
	A	B			3	A-E E-B	30	*26*	10.3	34
	A	C		52	4	A-E E-C	70	100	52	74
02:10	A	B	40		1	A-D D-B	*25. 3*	25	9.9	39
	A	B			2	A-D D-E E-B	*50. 5*	49	19.8	39

	A	B			3	A-E E-B	26.3	26	10.3	39
	A	C		52	4	A-E E-C	70	100	52	74
03:00	A	B	40		1	A-D D-B	20	22	911.1	56
	A	B			2	A-D D-E E-B	40	44	22.2	56
	A	B			3	A-E E-B	30	33	16.7	56
	A	C		52	4	A-E E-C	70	100	52	74

For rebalancing the traffic on existing MPLS paths, the traffic on each path *n* connecting same source and destination nodes is calculated as follows:

$$R_{Pn} = BW_{Pn} / \Sigma T_{S\text{-}D} \qquad \text{EQ3}$$

and the traffic on path Pn should be:

$$T'_{Pn} = T_{S\text{-}D} \times R_{Pn} \qquad \text{EQ4}$$

where the traffic from the source and destination nodes is denoted with $T_{S\text{-}D}$, R_{Pn} is the ratio of the traffic on path *n*, BW_{Pn} is the bandwidth of path *n*, T_{Pn} is the actual traffic on path *n*, and $\Sigma T_{S\text{-}D}$ is the sum of the traffic on all paths from source node *S* to destination node *D*.

Path occupancy PO is determined from the path bandwidth and the actual traffic:

$$PO = BW_{Pn} \times T_{Pn} / 100\% \qquad \text{EQ5}$$

For the example of Figure 17 and Table 1, path P1 carries a T_{Pn}=20 Mbit/s (first line of Table 1), and the sum $\Sigma T_{A\text{-}B}$ of all paths between nodes A and B is 20+40+30 = 90Mbit/s. The above EQ3 gives for the ratio of traffic on path P1 20/90=22%. Since the traffic between nodes A and B is 50Mbit/s, the actual traffic on path P1 is given by EQ4 as

50Mbit/s x 22% = 11Mbit/s.

When the T_{S-D} changes, as shown in the fifth line of table 1, the path occupancy for all three paths between node A and B change accordingly, from 56% to 94%. I.e.

$T'_{Pn} = T_{S-D} \times R_{Pn}$ =85Mbit/s x 22% = 18.9Mbit/s

PO = 20Mbit/s x 18.9Mbit/s/100%=94%

Figure 18A is a graph for comparing the ONC versus OSPF throughput;

In any network there is unused BW that is inaccessible because of redundant paths that are needed in case of link failure. These protection paths are not used, or are used for carrying lower priority traffic that is discarded in the case of a fault. Also, a path will never use 100% of the available bandwidth on the path. Furthermore, in most networks there is bandwidth that is reserved /committed to fulfil a QoS agreement.

The ONC allows the network to access this unused bandwidth by dynamically setting the MPLS paths, and establishing new connections at layer-1 or changing connections at layer-1 to adapt to changing traffic patterns.

Figure 18B is a graph showing the bandwidth usage for networks using ONC, versus ring and mesh protection. It is apparent the ONC ability to increase the volume of traffic that can be sent by using more of the available BW that cannot be accessed without an ONC. This holds true until the majority of the links on the network become highly connected. At this point, IP traffic following the open shortest path first (OSPF) protocol becomes more efficient than ONC-established MPLS paths. Although it would be extremely rare for a network to operate at such a congested level, should this occur, the ONC will prioritize OSPF.

The ONC also achieves efficiency through the reduction of BW required to be reserved for protection. The ONC has the ability to reserve resources for protection using both layer-1 and layer-3 resources, thereby reducing the amount of layer-1 resources needed to be reserved.

Each nodal ONC requires the pre-configured information listed in Table 2.

TABLE 2: Nodal Onc Configuration Information

Item	Learned From	Notes
ONC's own IP address	Pre-configuration during installation	From carrier's private IP address space.
Network Node identity	Pre-configuration during installation	Must be unique within the network.
Name Service address	Well-known	
Address of Management System	Nameserver using a well-known name	
Address of attached Service node	Nameserver using a well-known name	May not exist. From carrier's private IP address space.
Address of attached transport node	Nameserver using a well-known name. Note that the name and associated address may have to be manually configured in the nameserver.	May not exist. From carrier's private IP address space. This is the address that allowed the ONC to access the MIB (i.e. the SNMP address of the transport switch).
Capability of attached transport switch	Policy set manually	These include sufficient information for the nodal ONC to be able to create and initialize the appropriate objects.
Capability of attached core router	Policy set manually	

Inter-Node ONC Bandwidth

Bandwidth ID is required between ONCs for the exchange of information. This section attempts to estimate the required bandwidth.

The assumptions made for exchanges at the nodal-ONC layer once basic topology has been established are shown in Table 3.

TABLE 3: Bandwidth Sizing Assumptions

Item	Value	Notes
Nodes in area	64	
Unified Label-Switched Paths from my node	2048	64 nodes x 8 QoS x 8 routes
Length of statistics associated with a Unified Label-Switched path	8 bytes	
Number of transport switch connections	256	
Length of statistics associated with transport switch connection	4 bytes	Limited to status
Information exchange frequency	10 Hz	This is probably too high for initial implementations

The required bandwidth with these assumptions would be approximately:

$10 \times ((2048 \times 8 \times 8) + (256 \times 8 \times 8)) = (1.5) \times 10^6 \div 1.5$Mbit/s

Note that this assumes only one area for the ONC domain.

ONC Layering

The hierarchical model for the arrangement of ONCs in a large network is described next for an exemplary 64-node network, however, this hierarchical model will not be necessary.

The interaction of ONCs between domains can be though of as an exterior gateway protocol. Within an area, the optimization required of the ONC is greater than it is between areas. In a large network, areas may themselves be grouped hierarchically. The number of levels of hierarchy in any domain must be the same for all nodes. The highest level of hierarchy corresponds to the domain itself. As a minimum, a domain must contain at least one area.

Consider the network shown in Figure 19. Domain 1 includes five areas: **A**, **B**, **C**, **AB**, and **AB-C**. Area **A** is comprised of three nodes: **A1**, **A2**, **A3**. Area **B** is comprised of four nodes: **B1**, **B2**, **B3**, **B4**. Area **C** is comprised of one node: **C1**. Area **AB** is the combination of Areas **A**, **B**, and **C**. Domain 2 comprises one area: **D**.

Two independent dimensions of layers are present in the ONC:

- the ONC controls several layers of the telecommunications network: fibre, wavelength, label –switched path etc.

- the ONC itself is layered into hierarchical areas.

Consider now the network shown in Figure 20A. This represents a geographically-distributed network of twelve network nodes. The nodes have been grouped into areas in such a way that the boundaries make organizational sense to the carrier and logical sense as regards grouping by delay.

The nodes within Area A are connected by links with latency that does not affect the ONC's algorithms (below about 1ms or 300km). The nodes within Area-B and Area-D are also so connected. Area-AB is a combination of Area-A and Area-B. Area–AB has been designed to conform to the carrier's organization structure, as has Area-AB-C-D, which is a combination of Area-AB, Area-C and Area-D.

At least one ONC is associated with each node. In addition, there is an ONC associated with each area, arranged in a hierarchy as illustrated in Figure 20B. The roles of each of the ONC layers are as follows:

- sub-nodal ONCs: these are responsible for the collection of detailed information at a top-level (typically associated with a link or ULSP). They perform calculations to characterize the state of the low-level connections and pass this information to the nodal-level ONCs. In general, sub-nodal ONCs do not perform network rebalancing. Rules at this level are likely to be common (independent of network or node topology), remain stable and membership functions (the value of a parameter for a certain variable) are likely to change only rarely.

- nodal ONCs: these are responsible for analyzing the information received from the sub-nodal ONCs and making node-level decisions. A node-level decision involves rebalancing without modifying network topology. In general, rules at this level are largely common (network-independent), are stable and membership functions are likely to change

frequently to maintain discrimination-the effect of parameter changes on the rules-as conditions change.

- area ONCs: there are any number of these in the hierarchy. They are responsible for making area-wide decisions that could not be handled at the nodal level. Rules at this level are changed frequently to reflect network topology changes (as they embed the network topology) and membership functions are likely to change frequently to maintain discrimination.

Thus, the decision to route traffic from a failed link to another, spare, link on the same card might be taken at the sub-nodal level. A decision to adjust the fractions of traffic using two different links out of a node might be taken at the nodal level. The decision to route traffic through A1-A2-B1-B4 to offload A1-A3-B4 would require agreement at the Area-AB level without reference to Area-AB-C-D.

In summary, each ONC is responsible for summarizing the information at its level, making rebalancing decisions at its level and providing its parent with the summarized information.

Distribution of the code is, to some extent, a design decision and it is a architectural principle that it be flexible, but it is expected that the ONCs would be distributed as follows:

- sub-nodal ONCs: these reside as close to their interface points as possible, preferably on the linecards themselves, and if that is not possible, then on the nodal processor.

- nodal-ONCs: these run on the corresponding nodal processors.

- area-ONCs: these run in one of three ways:

- on all of the nodes within their area. This is likely to be true of at least the first few layers of the area hierarchy. The algorithms designed to allow such distributed ONCs to run are as yet undefined.

- on one of the nodes in the area. In this mode one ONC in each of the lower-layer groups is nominated (in PNNI-fashion) to act as the group's representative at the higher level. The algorithm for choosing the representative and reelecting the representative on failure is as yet undefined. This technique may be inappropriate in geographically small

network (where the number of layers may not warrant it) or in a global network (where the delays may prohibit it) but it may be required in continental-scale networks.

- on a centralized server that is not part of a node. In this mode, as in the previous mode described above, one ONC in each of the lower-layer groups is nominated to act as the group's representative at the higher level.

This technique is illustrated in Figure 21 where the network of Figure 20A has been mapped to a possible physical instantiation. In Figure 21, it has been decided that Area-A, Area-B, Area-C and Area-D should be run in distributed mode, but that Area-AB and Area-AB-C-D should be run in centralized mode. Thus, for example, the ONCs for sub-nodes A11, A12, ..., A1n, A1, and A all run on the processor associated with a single node, but there is only one ONC associated with Area-AB-C-D, which runs on a server in a centralized location. Area AB's ONC runs on A3.

ONC interfaces do not occur at all layers (see Table 4).

TABLE 4: ONC Interfaces

Interface	C1	C2	C5.1	C5.2	C7	C8	C9	M3
Sub-nodal		✓		✓		✓		✓
Nodal	✓		✓	✓			✓	✓
Area			✓	✓				✓
Top-Level are (domain)				✓	✓			✓

While the invention has been described with reference to particular example embodiments, further modifications and improvements which will occur to those skilled in the art, may be made within the purview of the appended claims, without departing from the scope of the invention in its broader aspect.

WE CLAIM:

1. A controller for an optical network comprising, at a traffic node:

a nodal network controller ONC;

a layer-3 interface between said nodal ONC and a service network element at said traffic node, for controlling a service node of said network to balance the traffic, adjust the bandwidth of a plurality of multiprotocol label-switching MPLS paths, and automatically set-up end-to-end MPLS paths;

a layer-1 interface between said nodal ONC and a transport network element at said traffic node, for defragmenting the optical network, to allow said optical network to make use of previously stranded resources; and

an in inter-layer interface to allow said transport network element to be used directly by the layer-3 to respond to congestion or increased demand at the layer-3.

2. A controller as claimed in claim 1, further comprising an intra-layer interface for communication of said nodal ONC with another nodal ONC.

3. A controller as claimed in claim 2, wherein said another nodal ONC is in same domain with said nodal ONC.

4. A controller as claimed in claim 2, wherein said another nodal ONC is in a different domain than said nodal ONC.

5. A controller as claimed in claim 1, further comprising a management interface for controlling said ONC by means of user-defined policies.

6. A controller as claimed in claim 5, wherein said management interface allows adjustment of policies for customizing the ONC's actions.

7. A controller as claimed in claim 5, wherein said management interface obtains records of audit trails and explanations of every action performed, and obtain recommendations from the ONC of potential actions to be performed.

8. A controller for an IP optical network comprising resource conservation means for maintaining the bandwidth between two traffic nodes to a level that is adjusted dynamically in accordance with the current traffic requirements.

9. A controller as claimed in claim 8, wherein said resource conservation means adjusts the bandwidth of a MPLS path according to the path utilization level.

10. A controller as claimed in claim 8, wherein said resource conservation means allocates additional bandwidth to a MPLS path whenever the traffic level on said path exceeds a high threshold.

11. A controller as claimed in claim 8, wherein said resource conservation means reduces the bandwidth allocated to a MPLS path whenever the traffic level on said path is below a low threshold.

12. A controller as claimed in claim 8, wherein said resource conservation means allocates additional bandwidth to a MPLS path whenever the traffic level on said path exceeds a high threshold.

13. A controller as claimed in claim 8, wherein said resource conservation means reduces bandwidth allocated to said MPLS path in decrements which are smaller than a previous increment.

14. A controller for an IP network comprising resource deployment means for redistributing resources between the MPLS paths.

15. A controller as claimed in claim 14, wherein said resource deployment means operates to first increase the bandwidth allocated to a MPLS path connecting two traffic nodes, whenever the traffic level on said path exceeds a preconfigured threshold.

16. A controller as claimed in claim 16, wherein said resource deployment means operates to create a new path between said two nodes if the bandwidth of said MPLS path cannot be increased.

17. A controller as claimed in claim 16, wherein said resource deployment means operates to create a new optical route by changing the topology of layer-1.

18. A layered network controller comprising a transport layer controller, a wavelength controller, and a service layer controller.

19. A network controller comprising a plurality of nodal controllers arranged in hierarchical areas.

20. A controller as claimed in claim 19, comprising a sub-nodal controller for collecting of detailed top-level information, determining the state of a low-level connections and passing this information to a nodal-level ONCs.

21. A controller as claimed in claim 19, comprising a nodal controller for analyzing the information received from said sub-nodal controller and making node-level decisions.

22. A controller as claimed in claim 21, wherein a node-level decision involves rebalancing said network without modifying network topology.

23. A controller as claimed in claim 19, comprising an area controller for making area-wide decisions that could not be handled at the nodal level.

ABSTRACT

A network controller which maximizes a carrier's revenue by distributing and optimizing network resources to increase the volume of traffic that can flow through the network is described. The network controller provides multi-terabit switching and routing capacities, dynamic network optimization, consolidation of multiple protocols into a unified network and is able to prioritize the highest revenue-generating traffic. The network controller works at three levels to optimize network resources. At the packet layer, the network controller balances the traffic, adjusts the bandwidth of multiprotocol label-switching (MPLS) paths, and automatically sets up end-to-end paths. At the optical layer, the network controller defragments the optical network, allowing the network to make use of previously stranded resources. Between the optical and packet layers, the network controller works to allow optical resources to be used directly by the packet layer to respond to congestion or increased demand at the packet layer.

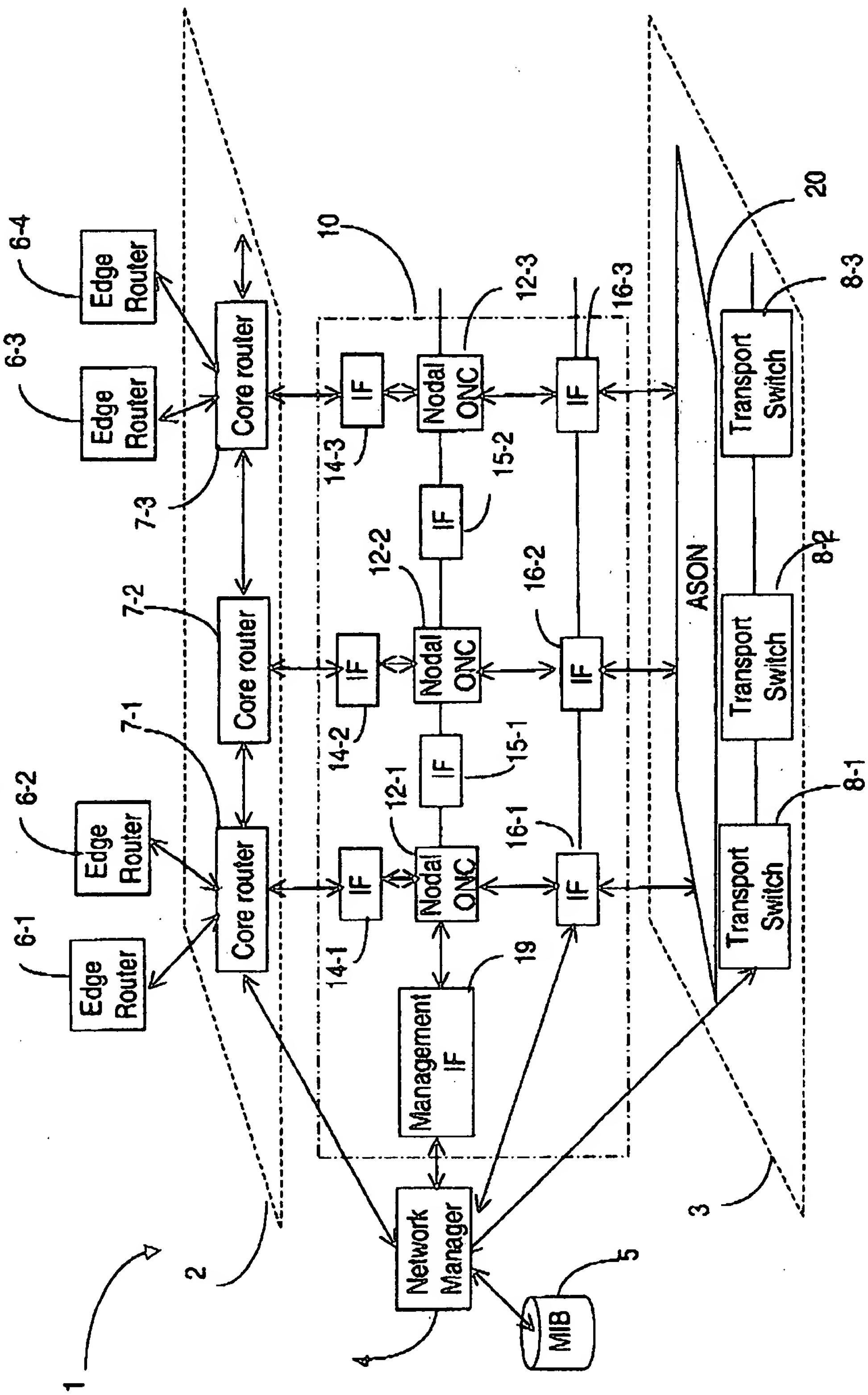

1
2
3
4
5
MIB
Network Manager
Management IF
19
Edge Router
6-1
6-2
6-3
6-4
Core router
7-1
7-2
7-3
IF
14-1
14-2
14-3
Nodal ONC
12-1
12-2
12-3
15-1
15-2
16-1
16-2
16-3
10
ASON
Transport Switch
8-1
8-2
8-3
20

FIGURE 14

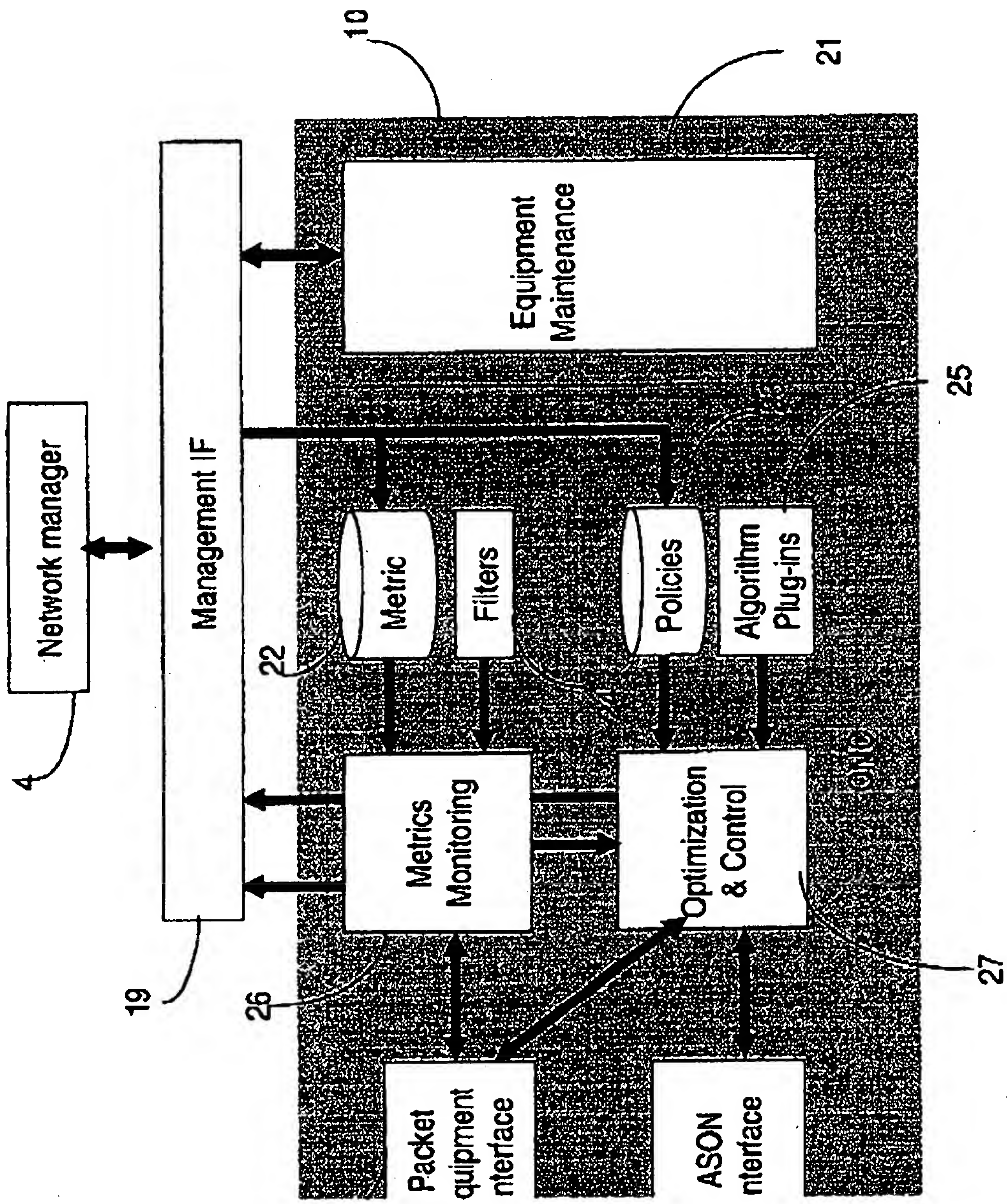

Network manager
Management IF
Equipment Maintenance
Metric
Filters
Policies
Algorithm Plug-ins
Metrics Monitoring
Optimization & Control
Packet quipment nterface
ASON nterface
4
10
19
21
22
25
26
27

FIGURE 15

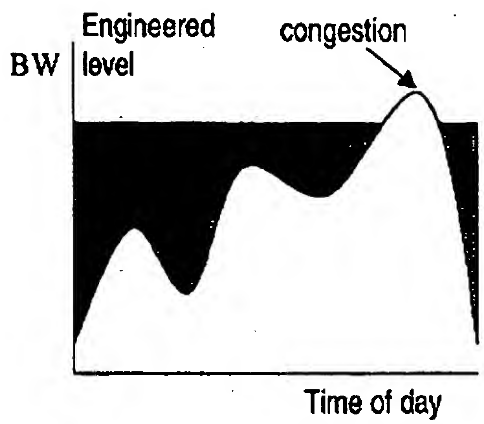


FIGURE 16A

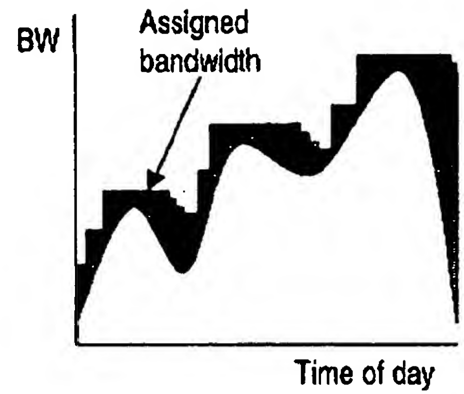


FIGURE 16B

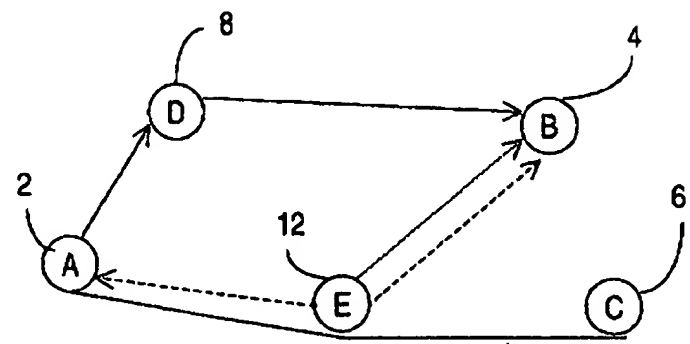


FIGURE 17

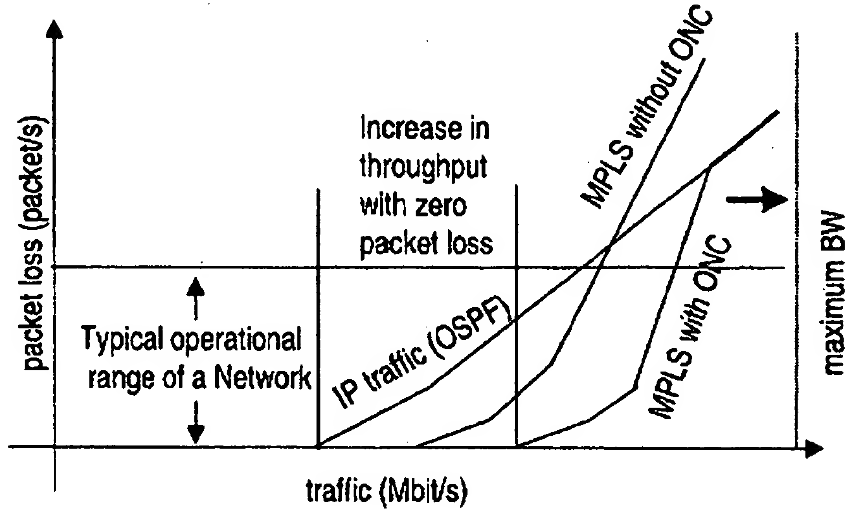


FIGURE 18A

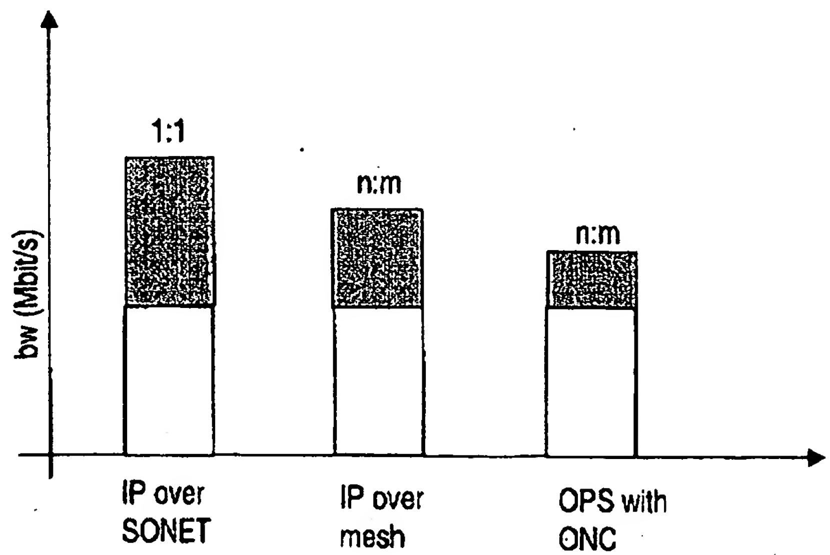


FIGURE 18B

60208946 .053100

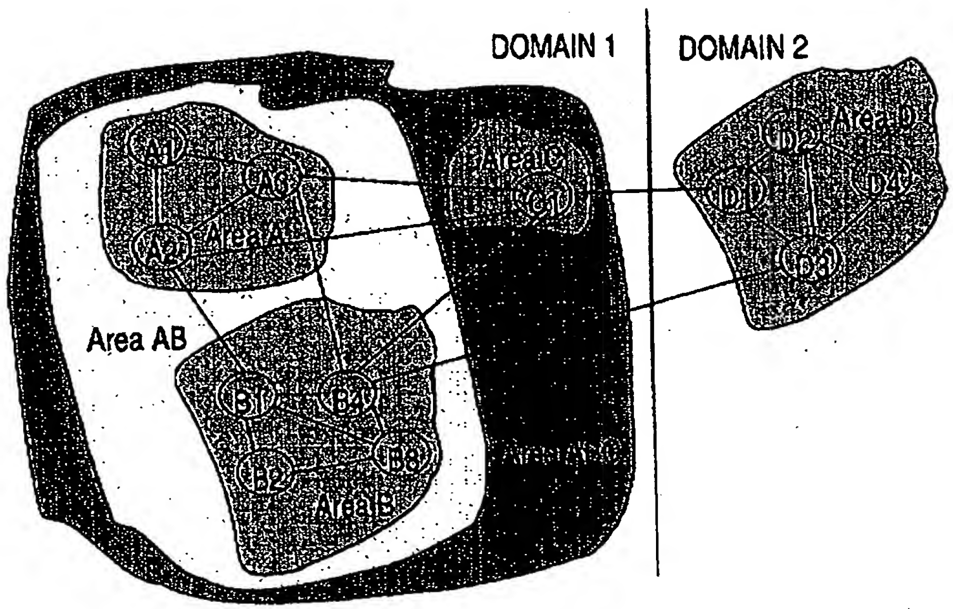

DOMAIN 1
DOMAIN 2
Area AB
A1
A2
AC
Area A
B1
B4
B2
B8
Area B
Area C
C1
D1
D2
D4
D3
Area D

FIGURE 19

FIGURE 20A

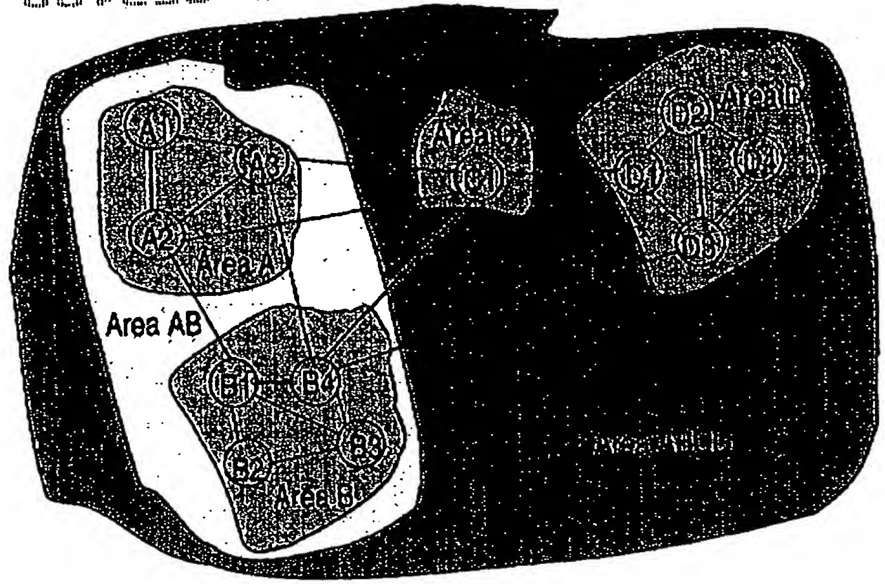


FIGURE 20B

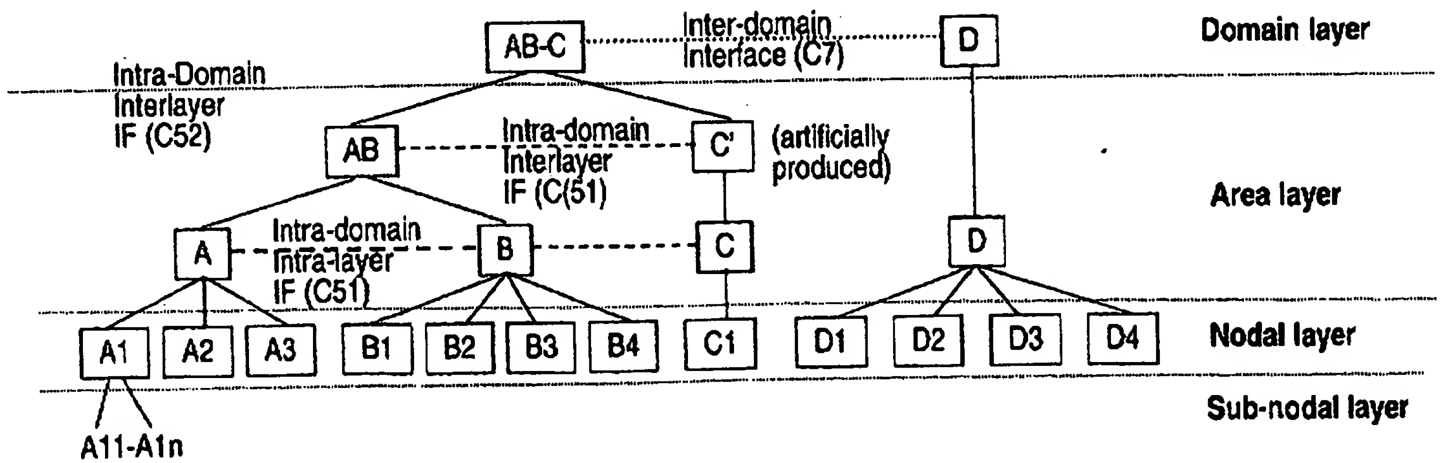
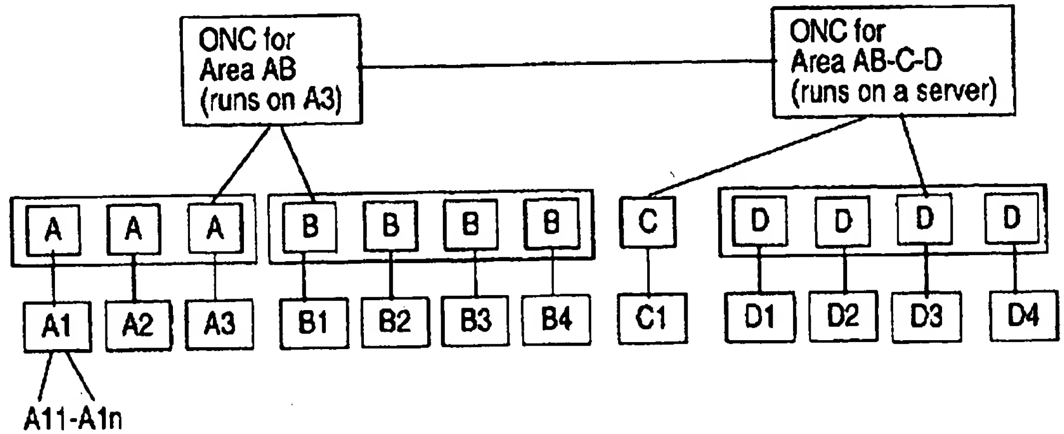


FIGURE 21


ONC for
Area AB
(runs on A3)
ONC for
Area AB-C-D
(runs on a server)
A
A
A
B
B
B
B
C
D
D
D
D
A1
A2
A3
B1
B2
B3
B4
C1
D1
D2
D3
D4
A11-A1n